

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 3.18182 Seconds
(without alignments)
182.496 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCGXXXCXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	40	78.4	400	1	VG78_HSVI1
2	38	74.5	186	1	DHML_METFL
3	38	74.5	187	1	DHML_METFL
4	37	72.5	12	1	CXL3_CONMR
5	36	70.6	11	1	CXL1_CONMR
6	36	70.6	13	1	CXL4_CONMR
7	36	70.6	61	1	CXL2_CONMR
8	35	68.6	55	1	MR4C_DROME
9	33	64.7	266	1	GNP1_GIALA
10	32	62.7	64	1	FER2_STRGO
11	32	62.7	75	1	MT_EISFO
12	32	62.7	83	1	HEPC_MOUSE
13	32	62.7	141	1	CT2B_MOUSE
14	32	62.7	1239	1	POLS_EEEV3
15	32	62.7	1240	1	POLS_EEEV3
16	32	62.7	1247	1	POLS_ANNVG
17	32	62.7	1253	1	POLS_SFV
18	32	62.7	1254	1	POLS_RRVN
19	32	62.7	1254	1	POLS_RRVV
20	31	60.8	65	1	TXM7_DENAN
21	31	60.8	72	1	YVAU_VACCC
22	31	60.8	84	1	HEPC_RAT
23	31	60.8	86	1	TXM2_DENAN
24	31	60.8	99	1	NTC1_HUMAN
25	31	60.8	155	1	NSU4_CATCO
26	31	60.8	198	1	DDL_STRMU
27	31	60.8	233	1	TPL_TREPA
28	31	60.8	239	1	NOO2_PARDE
29	31	60.8	245	1	IEO_NPVOP
30	31	60.8	261	1	IEO_NPVAC
31	31	60.8	299	1	AIC_RABIT
32	31	60.8	352	1	DMPP_PSESP
33	31	60.8	380	1	APFJ_HUMAN

RESULT 1

VG78_HSVI1
ID VG78_HSVI1 STANDARD: PRT: 400 AA.
AC Q00167;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical gene 78 zinc-binding protein.
GN 78.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Auburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M75136; AAA88180.1; -.
DR PIR: D36794; ZBBE14.
KW Hypothetical protein; Zinc; Zinc-finger.
SQ SEQUENCE 400 AA; 44167 MW; D941DB7738B95CA9 CRC64;
Query Match 78.4%; Score 40; DB 1; Length 400;
Best Local Similarity 50.0%; Pred No. 27;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 133 CCGATLCDSC 142

RESULT 2

DHML_METFL
ID DHML_METFL STANDARD: PRT: 186 AA.
AC Q50425;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methylamine dehydrogenase light chain precursor (EC 1.4.99.3) (MADH).
GN MAUA.
OS Methyllobacillus flagellatum.
OC Bacteria; Proteobacteria; Beta subdivision; Methylophilus group;
OC Methyllobacillus.
OX NCBI_TaxID=405;

O97666 macaca mula
Q18964 caenorhabdi
Q28661 ocyctolaqus
Q9qlp2 sus scrofa
P04070 homo sapien
P33587 mus musculu
P31394 rattus norv
Q9r283 rattus norv
Q9r244 mus musculu
P15423 human coron
P13897 western equ
P27285 sindbis vir

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=KT / ATCC 51484 / DSM 6875 / VKM B-1610;
RX  MEDLINE=95362696; PubMed=7635847;
RT  "Cloning, sequencing, and mutation of a gene for azurin in
RT  Methylobacillus flagellatum KT.";
RL  J. Bacteriol. 177:4575-4578(1995).
CC  -!- FUNCTION: METHYLAMINE DEHYDROGENASE CARRIES OUT THE OXIDATION OF
CC  METHYLAMINE. ELECTRONS ARE PASSED FROM METHYLAMINE DEHYDROGENASE
CC  TO AMICYANIN.
CC  -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + acceptor = RCHO + NH(3)
CC  + reduced acceptor.
CC  -!- COFACTOR: EACH OF THE SMALL SUBUNITS CONTAINS A COVALENTLY BOUND
CC  POO-LIKE COFACTOR CALLED TRYPTOPHAN TRYPTOPHYLAQUINONE (TTQ).
CC  -!- PATHWAY: Methyamine utilization.
CC  -!- SUBUNIT: TETRAMER OF TWO LIGHT AND TWO HEAVY CHAINS.
CC  -!- SUBCELLULAR LOCATION: Periplasmic.
CC  -!- SIMILARITY: HIGH, TO OTHER SPECIES MADH LIGHT CHAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch)
CC  -----
CC  EMBL; L37427; AAC41474.1; -.
CC  HSSP; P22619; 2BBK.
CC  InterPro; IPR004229; Me-amine-deh_L.
CC  Pfam; PF02975; Me-amine-deh_L; 1.
CC  Oxidoreductase; Electron transport; Periplasmic; TTQ; Signal.
CC  SIGNAL 1 56 POTENTIAL.
CC  CHAIN 57 186 METHYLAMINE DEHYDROGENASE LIGHT CHAIN.
CC  BINDING 112 112 TTQ (BY SIMILARITY).
CC  BINDING 163 163 TTQ (BY SIMILARITY).
CC  DISULFID 78 143 BY SIMILARITY.
CC  DISULFID 84 116 BY SIMILARITY.
CC  DISULFID 91 176 BY SIMILARITY.
CC  DISULFID 93 141 BY SIMILARITY.
CC  DISULFID 101 132 BY SIMILARITY.
CC  DISULFID 133 164 BY SIMILARITY.
CC  SEQUENCE 186 AA; 20174 MW; E94721FDA794B523 CRC64;

Query Match 74.5%; Score 38; DB 1; Length 186;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXC 13
Db 132 CCGKQTCGRC 141
||| | |
DEHML_METME STANDARD; PRT; 187 AA.
AC Q59543;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Methyamine dehydrogenase light chain precursor (BC 1.4.99.3) (MADH).
GN MAVA.
OS Methylophilus methylotrophus (Bacterium W3A1).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
CC Methylophilus.
CC NCBI_TaxID=17;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292427; PubMed=80211188;
RA Chistoserdov A.Y., McIntire W.S., Mathews F.S., Lidstrom M.E.;
RT "Organization of the methylamine utilization (mau) genes in
RT Methylophilus methylotrophus W3A1-NS.";

```

```

RL J. Bacteriol. 176:4073-4080(1994).
CC -!- FUNCTION: METHYLAMINE DEHYDROGENASE CARRIES OUT THE OXIDATION OF
CC METHYLAMINE. ELECTRONS ARE PASSED FROM METHYLAMINE DEHYDROGENASE
CC TO AMICYANIN.
CC -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + acceptor = RCHO + NH(3)
CC + reduced acceptor.
CC -!- COFACTOR: EACH OF THE SMALL SUBUNITS CONTAINS A COVALENTLY BOUND
CC POO-LIKE COFACTOR CALLED TRYPTOPHAN TRYPTOPHYLAQUINONE (TTQ).
CC -!- PATHWAY: Methyamine utilization.
CC -!- SUBUNIT: TETRAMER OF TWO LIGHT AND TWO HEAVY CHAINS.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- SIMILARITY: HIGH, TO OTHER SPECIES MADH LIGHT CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; L26407; AAB46951.1; -.
CC HSSP; P22619; 2BBK.
CC InterPro; IPR004229; Me-amine-deh_L.
CC Pfam; PF02975; Me-amine-deh_L; 1.
CC Oxidoreductase; Electron transport; Periplasmic; TTQ; Signal.
CC SIGNAL 1 57 POTENTIAL.
CC CHAIN 58 187 METHYLAMINE DEHYDROGENASE LIGHT CHAIN.
CC BINDING 113 113 TTQ (BY SIMILARITY).
CC BINDING 164 164 TTQ (BY SIMILARITY).
CC DISULFID 79 144 BY SIMILARITY.
CC DISULFID 85 117 BY SIMILARITY.
CC DISULFID 92 177 BY SIMILARITY.
CC DISULFID 94 142 BY SIMILARITY.
CC DISULFID 102 133 BY SIMILARITY.
CC DISULFID 134 165 BY SIMILARITY.
CC SEQUENCE 187 AA; 20237 MW; E33FEBAE30CC5CED CRC64;

Query Match 74.5%; Score 38; DB 1; Length 187;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXC 13
Db 133 CCGKQTCGRC 142
||| | |
CXL3_CONMR STANDARD; PRT; 12 AA.
AC P58809;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMRf.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
CC Neogastropoda; Conoidea; Conidae; Conus.
CC NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1262.77; MW_ERR=0.07; METHOD=Electrospray.

```

CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 3 12
FT DISULFID 4 9
FT MOD_RES 11 11 HYDROXYLATION.
SQ SEQUENCE 12 AA; 1251 MW; 277AAE2422D5A3C8 CRC64;

Query Match 72.5%; Score 37; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
| | | | |
Db 3 CCGYKLCPC 12

RESULT 5

CXL1_CONMR STANDARD; PRT; 11 AA.
ID CXL1_CONMR STANDARD; PRT; 11 AA.
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMrVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT DISULFID 3 8
FT MOD_RES 10 10 HYDROXYLATION.
SQ SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;

Query Match 70.6%; Score 36; DB 1; Length 11;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
| | | | |
Db 2 CCGYKLCPC 11

RESULT 6

CXL4_CONMR STANDARD; PRT; 13 AA.
ID CXL4_CONMR STANDARD; PRT; 13 AA.
AC P58810;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda/chi-conotoxin MrIB (Chi-MrIB).
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
RC TISSUE=Venom;

RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- PTM: Exists in two forms, due to cis-trans isomerization at His-
CC 11-Hyp-12.
CC -!- MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB: 1IEO; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT DISULFID 5 10
FT MOD_RES 12 12 HYDROXYLATION.
SQ SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 70.6%; Score 36; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
| | | | |
Db 4 CCGYKLCPC 13

RESULT 7

CXL2_CONMR STANDARD; PRT; 61 AA.
ID CXL2_CONMR STANDARD; PRT; 61 AA.
AC P58808;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMrVIB precursor (Chi-conotoxin MrIA) (Chi-MrIA)
DE (mr10a).
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 49-61, SYNTHESIS, AND MASS
RP SPECTROMETRY.
RC TISSUE=Venom duct, and Venom;
RX MEDLINE=20490660; PubMed=10900201;
RA McIntosh J.M., Corpuz G.O., Layer R.T., Garrett J.E., Wagstaff J.D.,
RA Bulaj G., Vyazovkina A., Yoshikami D., Cruz L.J., Olivera B.M.;
RT "Isolation and characterization of a novel conus peptide with apparent
RT antinociceptive activity.";
RL J. Biol. Chem. 275:32391-32397(2000).
RN [2]
RP SEQUENCE OF 49-61, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
RN [3]
RP SEQUENCE OF 49-61, SYNTHESIS, MASS SPECTROMETRY, AND STRUCTURE BY NMR.
RC TISSUE=Venom;
RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).

CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- PTM: Exists in two forms, due to cis-trans isomerization at His-59-Hyp-60.
 CC -!- MASS SPECTROMETRY: MW=1408.5; METHOD=Electrospray.
 CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
 KW Neurotoxin; Toxin; Hydroxylation; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 48
 FT PEPTIDE 49 61 LAMBDA-CONOTOXIN CMRVIB.
 FT DISULFID 52 61
 FT DISULFID 53 58
 FT MOD_RES 60 60 HYDROXYLATION.
 SQ SEQUENCE 61 AA; 6499 MW; F4DE5B5A97EB8DBA CRC64;

Query Match 70.6%; Score 36; DB 1; Length 61;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
 ||| | |
 DB 52 CCGYKLCGPC 61

RESULT 8
 M84C_DROME STANDARD; PRT; 55 AA.
 AC 001644; Q9VIA0;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Male specific sperm protein Mst84Dc.
 GN M84Dc OR CG17945.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=92102953; PubMed=1684716;
 RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
 RA Schaefer M.;
 RT "A cluster of four genes selectively expressed in the male germ line
 of *Drosophila melanogaster*.";
 RL Mech. Dev. 35:143-151(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 CC -!- TISSUE SPECIFICITY: TESTIS.
 CC -!- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
 CC -!- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
 MOTIFS.
 CC -!- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X67703; CAA47939.1; -;
 DR EMBL: AE003672; AAF54025.1; -;
 DR FlyBase: FBgn0004174; Mst84Dc.
 KW Spermatogenesis; Repeat; Multigene family.
 SQ SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

Query Match 68.6%; Score 35; Length 55;
 Best Local Similarity 50.0%; Pred. No. 38;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
 ||| | |
 DB 9 CCGYCCGPC 18

RESULT 9
 GNPL_GIALA STANDARD; PRT; 266 AA.
 AC 097439;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucosamine-6-phosphate isomerase 1 (EC 3.5.99.6) (Glucosamine-6-phosphate deaminase 1) (GNPDA 1) (GlcN6P deaminase 1).
 GN GP11.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR4 / Polish genotype;
 RA van Keulen H., Steimle P.A., Bulik D.A., Borowiak R.K., Jarroli E.L.;
 RT "Cloning of two Giardia glucosamine 6-phosphate isomerase genes only
 one of which is transcriptionally activated during encystation.";
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: D-glucosamine 6-phosphate + H(2)O = D-fructose
 6-phosphate + NH(3).
 CC -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
 ISOMERASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF050754; AAD02508.1; -
DR HSP; P09375; IDEA.
DR InterPro; IPR000457; Glucosamine_iso.
DR InterPro; IPR004547; NagB.
DR Pfam; PF01182; Glucosamine_iso; 1.
DR TIGRFAMs; TIGR00502; nagB; 1.
DR PROSITE; PS01161; GLC_GALNAC_ISOMERASE; 1.
KW Hydrolase; Carbohydrate metabolism.
FT ACT_SITE 67 BY SIMILARITY.
SQ SEQUENCE 266 AA; 29406 MW; 3EC301B8D6868386B CRC64;

Query Match 64.7%; Score 33; DB 1; Length 266;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
Db 255 CCGSTSC 261

RESULT 10
FER2_STRGO STANDARD; PRT; 64 AA.
AC P18325;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ferredoxin 2 (Fd-2).
GN SUBB.
OS Streptomyces griseolus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1909;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-43.
RC STRAIN=ATCC 11796;
RX MEDLINE=91105127; PubMed=1846297;
RA O'Keefe D.P., Gibson K.J., Emptage M.H., Lenstra R., Romesser J.A.,
RA Little P.J., Omer C.A.;
RT "Ferredoxins from two sulfonyleurea herbicide monooxygenase systems in
RT Streptomyces griseolus";
RL Biochemistry 30:447-455(1991).
CC -!- FUNCTION: ELECTRON TRANSPORT PROTEIN FOR THE CYTOCHROME P-450-SUO
CC SYSTEM.

CC -!- COFACTOR: BINDS A 3FE-4S CLUSTER.
CC -!- INDUCTION: By herbicides.
CC -!- SIMILARITY: >50% TO FERREDOXIN-1, AND TO B.STEATOTHERMOPHILUS,
CC C.THERMOACETICUM, AND D.AFRICANUS FERREDOXINS.
CC -!- SIMILARITY: TO THE N-TERMINAL OF R.FASCIANUS FERREDOXIN FAS2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M32239; AAA26826.1; -
DR PIR; B37915; B37915.
DR HSP; P00209; IFXD.
DR InterPro; IPR001080; 3FE4S_ferredoxin.
DR PRINTS; PR00352; 3FE4SFRDXIN.
KW Electron transport; Iron-sulfur; 3Fe-4S.
FT METAL 10 10 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
FT METAL 16 16 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
FT METAL 55 55 IRON-SULFUR (3FE-4S) (BY SIMILARITY).

SQ SEQUENCE 64 AA; 6647 MW; 3918DF9BDAAEE6D0C CRC64;
Query Match 62.7%; Score 32; DB 1; Length 64;
Best Local Similarity 57.1%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
Db 10 CCGAGSC 16

RESULT 11
MT_EISFO STANDARD; PRT; 75 AA.
ID MT_EISFO
AC P81695; O96822;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cadmium-metallothionein (MT) (Fragment).
GN MT.
OS Eisenia foetida (Common brandling worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
OX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE OF 1-41.
RA Gruber C., Dallinger R., Hunziker P., Gehrig P., Berger B.,
RA Stuerzenbaum S.;
RT "(Cd)-metallothionein from Eisenia foetida: evidence for
RT posttranslational processing to a functional one-domain protein.";
RL Submitted (FEB-1999) to the SWISS-PROT data bank.
RN [2]
RP SEQUENCE OF 2-75 FROM N.A.
RA Stuerzenbaum S.R., Guppy L., Morgan A.J., Kille P.;
RT "Metallothionein isolated from the common brandling worm Eisenia
RT foetida";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: POSTERIOR ALIMENTARY CANAL.

CC -!- INDUCTION: BY CADMIUM.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AJ236886; CAA15423.1; -
DR InterPro; IPR001396; Echnd_Methion4.
DR PRINTS; PR00873; MTECHINOIDEA.
KW Metal-binding; Metal-thiolate cluster; Cadmium.
FT NON_TER 1 1
FT CONFLICT 11 11 A -> Q (IN REF. 1).
SQ SEQUENCE 75 AA; 7411 MW; FDE2C308A606494B CRC64;

Query Match 62.7%; Score 32; DB 1; Length 75;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
Db 4 CCGKSTC 10

RESULT 12
HEPC_MOUSE STANDARD; PRT; 83 AA.
ID HEPC_MOUSE
AC Q9EQ21;
DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hepcidin precursor.
GN HAMP OR HEPC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX MEDLINE=21269329; PubMed=11113132;
RA Pigeon C., Ilyin G., Counselaud B., Leroyer P., Turlin B., Brissot P.,
Loreal O.;
RT "A new mouse liver-specific gene, encoding a protein homologous to
RT human antimicrobial peptide hepcidin, is overexpressed during iron
RT overload.";
RL J. Biol. Chem. 276:7811-7819(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP POSSIBLE FUNCTION IN IRON HOMEOSTASIS.
RX MEDLINE=21353006; PubMed=11447267;
RA Nicolas G., Bennoun M., Devaux I., Beaumont C., Grandchamp B.,
Kahn A., Vaulont S.;
RT "Lack of hepcidin gene expression and severe tissue iron overload in
RT upstream stimulatory factor 2 (USF2) knockout mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:8780-8785(2001).
CC -!- FUNCTION: Seems to act as a signaling molecule involved in the
CC maintenance of iron homeostasis. Seems to be required in
CC conjunction with HFE to regulate both intestinal iron absorption
CC and iron storage in macrophages. May also have antimicrobial
CC activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed in liver.
CC -!- SIMILARITY: BELONGS TO THE HEPcidin FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF297664; AAG49293.1; -.
CC EMBL; BC021587; AAH21587.1; -.
CC MGD; MGI:1933533; Hamp.
KW Antibiotic; Hormone; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 53 POTENTIAL.
FT CHAIN 59 83 HEPcidin.
FT DISULFD 65 71 POTENTIAL.
FT DISULFD 68 81 POTENTIAL.
FT DISULFD 69 80 POTENTIAL.
FT DISULFD 72 77 POTENTIAL.
SQ SEQUENCE 83 AA; 9352 MW; C76423EA46260B18 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 83;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CCGXXXCC 13
DB 71 CCNQCQCIC 80

RESULT 13
CT2B_MOUSE STANDARD; PRT; 141 AA.
ID CT2B_MOUSE
AC P12400;

DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CTLA-2-beta protein precursor (Fragment).
GN CTLA2B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/10-BR; TISSUE=T-cell;
RX MEDLINE=89276474; PubMed=2786470;
RA Denizot F., Brunet J.-F., Roustan P., Harper K., Suzan M.,
Luciani M.-F., Mattei M.-G., Golstein P.;
RT "Novel structures CTLA-2 alpha and CTLA-2 beta expressed in mouse
RT activated T cells and mast cells and homologous to cysteine
RT proteinase proteoglycans.";
RL Eur. J. Immunol. 19:631-635(1989).
CC -!- FUNCTION: NOT KNOWN, EXPRESSED IN ACTIVATED T-CELL.
CC -!- SIMILARITY: STRONG, TO THE PROPEPTIDE REGIONS OF CYSTEINE
CC PROTEASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15592; CAA33615.1; -.
CC PIR; S04925; S04925.
CC HSSP; P07711; ICTL.
CC MGD; MGI:88555; Ctl2b.
KW T-cell; Repeat; Signal.
FT NON_TER 1 1
FT SIGNAL <1 31 POTENTIAL.
FT CHAIN 32 141 CTLA-2-BETA PROTEIN.
FT DOMAIN 43 48 2 X 3 AA TANDEM REPEATS OF E-W-K.
FT REPEAT 43 45 1.
FT REPEAT 46 48 2.
SQ SEQUENCE 141 AA; 16248 MW; 232E2EA32E8F2AC0 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 141;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CCGXXXC 10
DB 107 CCGSSMC 113

RESULT 14
ID POLS_EEEV STANDARD; PRT; 1239 AA.
AC P08768;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
DE 6 kDa peptide; Spike glycoprotein E1].
OS Eastern equine encephalitis virus (Eastern equine encephalomyelitis
OS virus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11021;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=82V-2137;
RX MEDLINE=87282265; PubMed=2886548;
RA Chang G.-J., Trent D.W.;

RT "Nucleotide sequence of the genome region encoding the 26S mRNA of
 RT eastern equine encephalomyelitis virus and the deduced amino acid
 RT sequence of the viral structural proteins.";
 RL J. Gen. Virol. 68:2129-2142(1987).
 CC -!- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
 CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
 CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
 CC HEMAGGLUTININ.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05816; CAA29261.1; -;
 CC PIR; A26816; VHWVEE.
 CC HSP; P03315; IVCP.
 CC MEROPS; S03.001; -;
 CC InterPro; IPR002548; Alpha_E1-glycop.
 CC InterPro; IPR000936; Alpha_E2-glycop.
 CC InterPro; IPR002533; Alpha_E3-glycop.
 CC InterPro; IPR001836; Alpha_Core.
 CC InterPro; IPR000930; Togavirin.
 CC Pfam; PF00943; Alpha_E2-glycop; 1.
 CC Pfam; PF00944; Alpha_Core; 1.
 CC Pfam; PF01563; Alpha_E3-glycop; 1.
 CC Pfam; PF01589; Alpha_E1-glycop; 1.
 CC PRINTS; PR00798; TOGAVIRIN.
 CC Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
 KW Serine protease.
 FT CHAIN 1 259 COAT PROTEIN C.
 FT CHAIN 260 SPIKE GLYCOPROTEIN E3.
 FT CHAIN 323 SPIKE GLYCOPROTEIN E2.
 FT CHAIN 742 6 KDA PEPTIDE.
 FT CHAIN 743 SPIKE GLYCOPROTEIN E1.
 FT CHAIN 799 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 136 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 142 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 210 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT TRANSMEM 261 POTENTIAL.
 FT TRANSMEM 684 POTENTIAL.
 FT TRANSMEM 727 POTENTIAL.
 FT TRANSMEM 777 POTENTIAL.
 FT TRANSMEM 1211 POTENTIAL.
 FT CARBOHYD 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 932 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1239 AA; 137431 MW; 8C7664A405D2D41C CRC64;
 Query Match 62.7%; Score 32; DB 1; Length 1239;
 Best Local Similarity 57.1%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4 CCGXXXC 10
 Db 860 CCGATQC 866
 RESULT 15
 ID POLS_EEEV3 STANDARD; PRT; 1240 AA.
 AC P27284;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Structural polyprotein (p130) [Contains: Coat protein C (EC 3.4.21.-)
 DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
 DE 6 kDa peptide; Spike glycoprotein E1].
 OS Eastern equine encephalitis virus (strain va33[ten broeck]) (Eastern

OS equine encephalomyelitis virus).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 CC Alphavirus
 CC NCBI_TaxID=11022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91220727; PubMed=2024496;
 RA Weaver S.C., Scott T.W., Rico-Hesse R.;
 RT "Molecular evolution of eastern equine encephalomyelitis virus in
 RT North America";
 RL Virology 182:774-784(1991).
 CC -!- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
 CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -!- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
 CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
 CC HEMAGGLUTININ.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M69094; AAA42980.1; -;
 CC PIR; A39992; VHWVEV.
 CC HSP; P03315; IVCP.
 CC MEROPS; S03.001; -;
 CC InterPro; IPR002548; Alpha_E1-glycop.
 CC InterPro; IPR000936; Alpha_E2-glycop.
 CC InterPro; IPR002533; Alpha_E3-glycop.
 CC InterPro; IPR001836; Alpha_Core.
 CC InterPro; IPR000930; Togavirin.
 CC Pfam; PF00943; Alpha_E2-glycop; 1.
 CC Pfam; PF00944; Alpha_Core; 1.
 CC Pfam; PF01563; Alpha_E3-glycop; 1.
 CC Pfam; PF01589; Alpha_E1-glycop; 1.
 CC PRINTS; PR00798; TOGAVIRIN.
 CC Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
 KW Serine protease.
 FT CHAIN 1 260 COAT PROTEIN C.
 FT CHAIN 261 SPIKE GLYCOPROTEIN E3.
 FT CHAIN 323 SPIKE GLYCOPROTEIN E2.
 FT CHAIN 743 6 KDA PEPTIDE.
 FT CHAIN 744 SPIKE GLYCOPROTEIN E1.
 FT CHAIN 800 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 137 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 143 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 211 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT TRANSMEM 259 POTENTIAL.
 FT TRANSMEM 695 POTENTIAL.
 FT TRANSMEM 722 POTENTIAL.
 FT TRANSMEM 781 POTENTIAL.
 FT TRANSMEM 1212 POTENTIAL.
 FT CARBOHYD 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 834 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 933 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1240 AA; 137290 MW; AEBEB1599D083045 CRC64;
 Query Match 62.7%; Score 32; DB 1; Length 1240;
 Best Local Similarity 57.1%; Pred. No. 5.3e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 4 CCGXXXC 10
 Db 861 CCGATQC 867

Search completed: January 29, 2003, 09:29:22
 Job time : 5.18182 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:57 ; Search time 5.88636 Seconds
(without alignments)
228.644 Million cell updates/sec

Title: US-09-580-201A-1

Perfect score: 51

Sequence: 1 XXXCGXXXCCXXC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	78.4	400	1 ZBBI4	44.1K zinc-binding
2	39	76.5	286	2 S61199	hypothetical prote
3	39	76.5	1877	2 T21861	hypothetical prote
4	38	74.5	187	2 T10073	amine dehydrogenas
5	38	74.5	258	2 T30368	probable immediate
6	38	74.5	328	2 E97129	probable flavodoxi
7	38	74.5	948	2 A57640	retinoblastoma bin
8	38	74.5	1560	2 T42727	proliferation pote
9	37	72.5	186	2 A45910	ultra-high-sulfur
10	37	72.5	196	2 T25136	hypothetical prote
11	37	72.5	264	2 F97086	anaerobic sulfite
12	37	72.5	312	2 T25135	hypothetical prote
13	37	72.5	550	2 G85436	hypothetical prote
14	37	72.5	693	2 S49228	sodium-dependent p
15	36	70.6	404	2 T02396	hypothetical prote
16	36	70.6	514	2 T10559	hypothetical prote
17	36	70.6	670	2 T49510	fibroin-3 related
18	36	70.6	756	2 D96527	protein F27J15.24
19	35	68.6	55	2 S25774	testis-specific pr
20	35	68.6	127	2 T34270	hypothetical prote
21	35	68.6	428	2 T46025	hypothetical prote
22	34	66.7	215	2 T16542	hypothetical prote
23	33	64.7	412	2 F86296	hypothetical prote
24	32	62.7	64	2 B37915	ferredoxin [3Fe-4S
25	32	62.7	141	2 S04925	CTLA-2-beta protei
26	32	62.7	194	2 T25548	hypothetical prote
27	32	62.7	406	2 T04817	hypothetical prote
28	32	62.7	422	2 D86446	hypothetical prote
29	32	62.7	547	2 T30269	hypothetical prote

30	32	62.7	1031	2 T43458	hypothetical prote
31	32	62.7	1145	2 S37136	structural polypro
32	32	62.7	1239	1 VHWEE	structural polypro
33	32	62.7	1240	1 VHWEE	structural polypro
34	32	62.7	1241	2 S26373	genome polyprotein
35	32	62.7	1242	2 S72350	structural polypro
36	32	62.7	1242	2 A56605	structural polypro
37	32	62.7	1247	1 VHWV2	structural polypro
38	32	62.7	1253	1 VHWV	structural polypro
39	32	62.7	1254	1 VHWVRA	structural polypro
40	31	60.8	38	2 T12331	metallothionein -
41	31	60.8	72	2 H42525	A-ORF-U protein -
42	31	60.8	86	2 A37910	muscarinic toxin 2
43	31	60.8	98	2 JC5147	tachycitin precurs
44	31	60.8	106	2 S77325	ferredoxin [2Fe-2S
45	31	60.8	154	2 E87530	isoquinoline 1-oxi

ALIGNMENTS

RESULT 1

ZBBI4

44.1K zinc-binding protein - ictalurid herpesvirus 1 (strain auburn 1)

C:Species: ictalurid herpesvirus 1

A:Note: host Ictalurus punctatus (channel catfish)

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: D36794

R.Davison, A.J.

submitted to GenBank, January 1992

A:Description: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A36804

A:Accession: D36794

A:Molecule type: DNA

A:Residues: 1-400 <DAV>

A:Cross-references: GB:M75136; NID:g331209; PIDN:AAA88180.1; PID:g331286

R.Davison, A.J.

Virology 186, 9-14, 1992

A:Title: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A39447; MUID:92087490; PMID:1727613

A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 78

C:Superfamily: ictalurid herpesvirus 44.1K zinc binding protein

C:Keywords: zinc finger

Query Match 78.4%; Score 40; DB 1; Length 400;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
| | | | |
Db 133 CCGATLCDSC 142

RESULT 2

S61199

hypothetical protein YDR313c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D9740.20

C:Species: Saccharomyces cerevisiae

C>Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002

C:Accession: S61199

R.Ding, H.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of S. cerevisiae cosmid 9740.

A:Reference number: S61160

A:Accession: S61199

A:Molecule type: DNA

A:Residues: 1-286 <DIN>

A:Cross-references: EMBL:U28374; NID:g849207; PID:g849227; GSPDB:GN00004; MIPS:YDR313

C:Genetics:

A:Gene: SGD:PIB1; MIPS:YDR313C

A:Cross-references: SGD:S0002721
A:Map position: 4R

Query Match 76.5%; Score 39; DB 2; Length 286;
Best Local Similarity 50.0%; Pred. No. 1.e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| | |
Db 41 CCGRIFCSC 50

RESULT 3

T21861
hypothetical protein F36F2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21861

R:Cottage, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19479

A:Accession: T21861

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1877 <WIL>

A:Cross-references: EMBL:Z81532; PIDN:CAB04326.1; GSPDB:GN00019; CESP:F36F2.3

A:Experimental source: clone F36F2

C:Genetics:

A:Gene: CESP:F36F2.3

A:Map position: 1

A:Introns: 49/3; 86/2; 112/3; 139/2; 235/3; 284/3; 436/3; 507/2; 566/2; 678/2; 1000/3; 1

Query Match 76.5%; Score 39; DB 2; Length 1877;

Best Local Similarity 50.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| | |
Db 320 CCGSYCADC 329

RESULT 4

T10073

amine dehydrogenase (EC 1.4.99.3) light chain - Methylophilus methylotrophus

C:Species: Methylophilus methylotrophus

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000

C:Accession: T10073

R:Chistoserdov, A.Y.; McIntire, W.S.; Mathews, F.S.; Lidstrom, M.E.

J. Bacteriol. 176, 4073-4080, 1994

A:Title: Organization of the methylamine utilization (mau) genes in Methylophilus methyl

A:Reference number: Z16936; MUID:94292427; PMID:8021188

A:Accession: T10073

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-187 <CHI>

A:Cross-references: EMBL:L26407; NID:g561931; PIDN:AAB46951.1; PID:g556334

A:Experimental source: strain W3A1

C:Genetics:

A:Gene: mauA

C:Complex: heterotetramer; two large and two small chains

C:Superfamily: amine dehydrogenase light chain

C:Keywords: oxidoreductase

Query Match 74.5%; Score 38; DB 2; Length 187;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| | |
Db 133 CCGKQTCGR 142

RESULT 5

T30368

probable immediate-early transactivator 0 - Lymantria dispar nuclear polyhedrosis vir

C:Species: Lymantria dispar nuclear polyhedrosis virus, LDMNPV

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T30368

R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R

Virolgy 253, 17-34, 1999

A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri

A:Reference number: Z20836; MUID:99124785; PMID:9887315

A:Accession: T30368

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-258 <KUZ>

A:Cross-references: EMBL:AF081810; PIDN:RAC70206.1

C:Keywords: immediate-early protein

Query Match 74.5%; Score 38; DB 2; Length 258;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| | |
Db 209 CCGYRVCNAC 218

RESULT 6

E97129

probable flavodoxin oxidoreductase CAC1860 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: E97129

R:Nolling, J.; Berton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: E97129

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-328 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79824.1; PID:g15024837; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1860

Query Match 74.5%; Score 38; DB 2; Length 328;

Best Local Similarity 50.0%; Pred. No. 1.4e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| | |
Db 290 CCGEGICGAC 299

RESULT 7

A57640

retinoblastoma binding protein RBQ-1 - human

C:Species: Homo sapiens (man)

C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 17-Mar-2000

C:Accession: A57640

R:Sakai, Y.; Saijo, M.; Coelho, K.; Kishino, T.; Niikawa, N.; Taya, Y.

Genomics 30, 98-101, 1995

A:Title: cDNA sequence and chromosomal localization of a novel human protein, RBQ-1 (

A:Reference number: A57640; MUID:96129310; PMID:8595913

A:Accession: A57640

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-948 <SAK>

A:Cross-references: GB:X85133; NID:g728590; PIDN:CAA59445.1; PID:g755748

C:Genetics:

A:Gene: GDB:RBBP6

A:Cross-references: GDB:626076

A:Map position: 16p12-16p11.2

C:Superfamily: RING finger homology
C:Keywords: tandem repeat; zinc
F:79-129/Domain: RING finger homology <RNG>

Query Match 74.5%; Score 38; DB 2; Length 948;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |

Db 98 CCGNSYCDEC 107

RESULT 8

T42727

proliferation potential-related protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T42727

R:Witte, M.M.; Scott, R.E.

submitted to the EMBL Data Library, November 1998

A:Reference number: Z22246

A:Accession: T42727

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1560 <WIT>

A:Cross-references: EMBL:U03913; NID:g3858884; PID:g3858885; PIDN:AAC72432.1

A:Experimental source: strain Balb/C

C:Genetics:

A:Gene: P2P-R

C:Function:

A:Description: involved in hnRNP association and Rb1 binding

C:Superfamily: RING finger homology

F:57-107/Domain: RING finger homology <RRN>

Query Match 74.5%; Score 38; DB 2; Length 1560;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

||| | |

Db 76 CCGNSYCDEC 85

RESULT 9

A45910

ultra-high-sulfur keratin - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 24-Sep-1999

C:Accession: A45910

R:McNab, A.R.; Wood, L.; Theriault, N.; Gierman, T.; Vogeli, G.

J. Invest. Dermatol. 92, 263-266, 1989

A>Title: An ultra-high sulfur keratin gene is expressed specifically during hair growth.

A:Reference number: A45910; MUID:89140394; PMID:2465353

A:Accession: A45910

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-186 <MCN>

A:Cross-references: GB:M27685; NID:g341749; PIDN:AAA81560.1; PID:g1066818

C:Superfamily: ultra-high-sulfur keratin

Query Match 72.5%; Score 37; DB 2; Length 186;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

||| | |

Db 42 CCGSCCQC 51

RESULT 10

T25136

hypothetical protein T22H2.5a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T25136

R:Lennard, N.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19985

A:Accession: T25136

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-196 <WIL>

A:Cross-references: EMBL:Z81595; PIDN:CAB04755.1; GSPDB:GN00019; CESP:T22H2.5a

A:Experimental source: clone T22H2

C:Genetics:

A:Gene: CESP:T22H2.5a

A:Map position: 1

A:Introns: 25/3; 126/3; 161/1; 180/3; 196/1

Query Match 72.5%; Score 37; DB 2; Length 196;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

||| | |

Db 58 CCGCLACIGC 67

RESULT 11

F97086

anaerobic sulfite reductase, B chain [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 01-Mar-2002

C:Accession: F97086

R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L.

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: F97086

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-264 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK79481.1; PID:gl5024461; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1514

C:Superfamily: Pyrococcus furiosus cytochrome-c3 hydrogenase gamma chain

Query Match 72.5%; Score 37; DB 2; Length 264;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

||| | |

Db 231 CCGIGKCGHC 240

RESULT 12

T25135

hypothetical protein T22H2.5b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T25135

R:Lennard, N.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19985

A:Accession: T25135

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-312 <WIL>

A:Cross-references: EMBL:Z81595; PIDN:CAB04754.1; GSPDB:GN00019; CESP:T22H2.5b

A:Experimental source: clone T22H2

C:Genetics:

A:Gene: CESP:T22H2.5b

A:Map position: 1
A:Introns: 135/3; 236/3; 271/1; 290/3; 306/1
C:Superfamily: Caenorhabditis elegans hypothetical protein K08D10.7

Query Match 72.5%; Score 37; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 168 CCGCLACIGC 177

RESULT 13
G85436
hypothetical protein At2g436980 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: G85436
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: G85436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <STO>
A:Cross-references: GB:NC_001268; NID:g7270647; PIDN:CAB80364.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g36980
A:Map position: 4

Query Match 72.5%; Score 37; DB 2; Length 550;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 499 CCGSFFCCLC 508

RESULT 14
S49228
sodium-dependent phosphate transporter - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 05-Nov-1999
C:Accession: S68972; S49228
R:Helps, C.; Murer, H.; McGivan, J.
Eur. J. Biochem. 228, 927-930, 1995
A:Title: Cloning, sequence analysis and expression of the cDNA encoding a sodium-depende
A:Reference number: S68972; MUID:95255303; PMID:7737195
A:Accession: S68972
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-693 <HE2>
A:Cross-references: EMBL:X81699; NID:g547483; PIDN:CAA57345.1; PID:g547484

Query Match 72.5%; Score 37; DB 2; Length 693;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 628 CCGLCGCKC 637

RESULT 15
T02396
hypothetical protein At2g44410 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F411.22
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02396; B84878

R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K
submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
A:Reference number: Z14667
A:Accession: T02396
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-404 <ROU>
A:Cross-references: EMBL:AC004521; NID:g3128166; PID:g3128183
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, R.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84878
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-404 <STO>
A:Cross-references: GB:AE002093; NID:g3128183; PIDN:AAC16087.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g44410; F411.22
A:Map position: 2
C:Superfamily: RING finger homology
F:112-162/Domain: RING finger homology <RRN>

Query Match 70.6%; Score 36; DB 2; Length 404;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 130 CCGHFCWGC 139

Search completed: January 29, 2003, 09:32:58
Job time : 7.88636 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 14 Seconds
(without alignments)
133.251 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCGXXCXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	78.4	210	AAW85717	Novel protein (ClO
2	40	78.4	300	ABO8606	TAP protein SEQ ID
3	40	78.4	1679	AAU07343	I-aminocyclopropan
4	39	76.5	31	AAU23346	Peptide sequence u
5	39	76.5	31	AA19699	4-Coumarate CoA li
6	39	76.5	31	AAE05833	Liquidambar sp. an
7	39	76.5	169	AAE43505	Human cancer assoc
8	39	76.5	246	AAW53007	Mus musculus i-mfa
9	39	76.5	349	ABO8605	Hirudin protein SE
10	39	76.5	919	AAG84970	Shrimp white spot

11	39	76.5	1231	22	ABB59738	Drosophila melanog
12	39	76.5	1404	23	AAE20581	Proliferation pote
13	38	74.5	80	22	AAU43831	Protonibacterium
14	38	74.5	153	21	AAU70234	Human RNA-associat
15	38	74.5	237	22	ABB69415	Drosophila melanog
16	38	74.5	973	22	AAU41045	Human polypeptide
17	38	74.5	1023	22	AAU11436	D. discoideum ster
18	38	74.5	1278	22	AAU39259	Human polypeptide
19	38	74.5	1329	23	AAU91279	Human NOV3a protei
20	37	72.5	166	22	AAU41699	Protonibacterium
21	37	72.5	225	21	AAU36084	Neospora caninum N
22	37	72.5	250	22	AAU79002	Human protein SEQ
23	37	72.5	301	22	AAU79986	Human protein SEQ
24	37	72.5	369	22	ABG15487	Novel human diagno
25	37	72.5	720	23	AAO14994	Laminin-related pr
26	36	70.6	12	21	AAU08014	Generic formula fo
27	36	70.6	12	21	AAU08015	Generic formula fo
28	36	70.6	12	21	AAU08018	Amino acid sequenc
29	36	70.6	12	21	AAU08019	Amino acid sequenc
30	36	70.6	13	21	AAU08012	Generic formula fo
31	36	70.6	13	21	AAU08013	Generic formula fo
32	36	70.6	13	21	AAU08017	Amino acid sequenc
33	36	70.6	13	21	AAU92229	Chi-conotoxin pept
34	36	70.6	13	21	AAU92230	Chi-conotoxin pept
35	36	70.6	21	19	AAU69478	Addition sequenc
36	36	70.6	61	21	AAU08016	Amino acid sequenc
37	36	70.6	61	21	AAU92231	Chi-conotoxin, chi
38	35	68.6	55	22	ABB66559	Drosophila melanog
39	34	66.7	149	21	AAG34801	Arabiopsis thalia
40	34	66.7	377	23	AAU75543	Human B7-like prot
41	34	66.7	382	23	AAU75540	Human B7-like prot
42	34	66.7	386	23	AAU75541	Human B7-like prot
43	34	66.7	386	23	AAU75542	Human B7-like prot
44	34	66.7	463	22	ABG28169	Novel human diagno
45	33	64.7	51	20	AAU30722	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAW85717
ID AAW85717 standard; Protein; 210 AA.
XX
AC AAW85717;
XX
DT 27-SEP-1999 (first entry)
XX
DE Novel protein (Clone AC41_1).
XX
KW Polynucleotide; protein; nutrition; cytokine; cell proliferation;
KW cell differentiation; immunostimulation; immunosuppression;
KW haematopoiesis regulation; tissue growth; activin; inhibin;
KW chemotaxis; chemokinesis; haemostasis; thrombolysis; receptor;
KW ligand; anti-inflammatory; tumour suppression; gene therapy.
XX
OS Homo sapiens.
XX
PN WO920644-A1.
XX
PD 29-APR-1999.
XX
PF 16-OCT-1998; 98WO-US22034.
XX
PR 18-OCT-1997; 97US-0955557.
XX
PA (GEM) GENETICS INST INC.
XX
PI Agostino MJ, Bowman MR, Evans C, Jacobs K, Lavallie ER;
PI McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
PI WPI; 1999-288272/24.
DR N-PSDB; AAX08682.

XX New polynucleotides encoding secreted human proteins
 PT
 PS Claim 8; Page 100; 136pp; English.
 XX
 CC The new human secreted proteins are encoded by polynucleotides
 CC obtained from human placenta, adult testes, fetal kidney, fetal
 CC brain, adult brain, adult brain and adult blood cDNA libraries.
 CC The polynucleotides and proteins are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals. Suggested
 CC activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, and tumour inhibition
 CC activity. The polynucleotides are also stated to be useful for gene
 CC therapy. The sequences identified by a secretory leader
 CC sequence motif in the polynucleotide and it is thought that the
 CC encoded proteins have biological activity by virtue of their secreted
 CC nature. This polypeptide was encoded by a clone designated AC41_1.
 CC (See AAX08682).
 XX

SQ Sequence 210 AA;

Query Match 78.4%; Score 40; DB 20; Length 210;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXXXC 13

||| | |

Db 13 CCGSAGCSLC 22

NOI

RESULT 2

ABB08606

ID ABB08606 standard; Protein; 300 AA.

XX

AC ABB08606;

XX 10-APR-2002 (first entry)

XX TAP protein SEQ ID NO 17.

XX Fusion protein; hirudin; Hir; tick anticoagulant protein; TAP;
 KW anticoagulant; blood; thrombin; factor Xa.

XX Ornithodoros moubata.

XX WO200204486-A2.

XX 17-JAN-2002.

XX 27-JUN-2001; 2001WO-EP07333.

XX 07-JUL-2000; 2000DE-1033195.

XX (AVET) AVENTIS PHARMA DEUT GMBH.

XX Habermann P;

XX WPI; 2002-154918/20.

XX New fusion protein of hirudin and tick anticoagulant protein, useful as
 PT an anticoagulant -
 PT
 PS Claim 4; Page 31-32; 36pp; German.
 XX
 CC The invention relates to a bifunctional fusion protein comprising
 CC hirudin or its variant and tick anticoagulant protein (TAP) or its
 CC variant. The fusion proteins have anticoagulant activity and are used to

CC inhibit coagulation of blood by acting to inhibit thrombin and inhibit
 CC factor Xa. The present sequence is that of the TAP protein, the first 60
 CC amino acids of which are given in figure 5.

SQ Sequence 300 AA;

Query Match 78.4%; Score 40; DB 23; Length 300;

Best Local Similarity 50.0%; Pred. No. 1.8e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXXXC 13

||| | |

Db 196 CCGTGACTGC 205

RESULT 3

AAU07343

ID AAU07343 standard; Protein; 1679 AA.

XX

AC AAU07343;

XX 04-DEC-2001 (first entry)

XX 1-aminocyclopropane carboxylate (ACPC) synthase #12.

KW 1-aminocyclopropane carboxylate synthase; ACPC synthase; brain trauma;
 KW excitatory neurotransmission; neurodegeneration; stroke; nerve damage;
 KW neurodegenerative disease; Alzheimer's disease; depression; epilepsy;
 KW alcohol abuse; cognitive function; memory; learning impairment; human.
 XX

OS Homo sapiens.

XX WO200168879-A2.

XX 20-SEP-2001.

XX 14-MAR-2001; 2001WO-EP02857.

XX 14-MAR-2000; 2000US-0189086.

XX 05-APR-2000; 2000US-0194702.

XX (FARB) BAYER AG.

XX Ramakrishnan S;

XX WPI; 2001-550286/61.

XX Isolated polynucleotide encoding a human
 PT 1-aminocyclopropane-carboxylate (ACPC) synthase, useful for treating
 PT brain trauma and neurodegenerative disease (e.g. Alzheimer's disease,
 PT depression, epilepsy) -
 XX
 PS Claim 1; Page 231-237; 242pp; English.

XX The invention relates to reagents and methods for regulating excitatory
 CC neurotransmission, and to prevent neurodegeneration. The method involves
 CC the use of an expression vector or a reagent that modulates the activity
 CC of a 1-aminocyclopropane-carboxylate (ACPC) synthase polypeptide. The
 CC reagent is useful for modulating the activity of an ACPC synthase in a
 CC disease such as stroke, a nerve damage or a neurodegenerative disease.
 CC The ACPC synthase polypeptide, polynucleotides and modulators are also
 CC useful for treating brain trauma and neurodegenerative disease (e.g.
 CC Alzheimer's disease, depression, epilepsy). The ACPC synthase modulators
 CC are also useful for treating alcohol abuse and improve cognitive function
 CC and memory of patients with learning impairment. The present function
 CC represents the amino acid sequence of human 1-aminocyclopropane-
 CC carboxylate (ACPC) synthase #12, used in the method of the invention.

SQ Sequence 1679 AA;

Query Match 78.4%; Score 40; DB 22; Length 1679;

Best Local Similarity 50.0%; Pred. No. 4.6e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13
 III I I
 Db 1112 CCGTGGCTTC 1121

RESULT 4

AA23346
 ID AAY23346 standard; Peptide; 31 AA.

AC AAY23346;

DT 01-SEP-1999 (first entry)

XX Peptide sequence used to derive a primer for CL gene isolation.

XX Coumarate CoA ligase; syringyl lignin; angiosperm; gymnosperm;
 KW loblolly pine; Pinus taeda; lignin; woody plant; paper manufacture;
 KW delignification; pulp mill.

XX Unidentified.

OS WO9931243-Al.

PN 24-JUN-1999.

PD 16-DEC-1998; 98WO-US26784.

PF 16-DEC-1997; 97US-0991677.

PR (INTO) INT PAPER CO.

PI Carraway DT, Chiang VL, Smeltzer RH;

DR WPI; 1999-405034/34.

XX New DNA sequences encoding angiosperm enzymes involved in
 PT biosynthesis of syringyl lignin

PS Example 1; Page 11; 83pp; English.

CC AAY23346-48 represent peptide sequences used to derived degenerate
 CC primers for coumarate CoA ligase (CL) gene isolation. CL is
 CC involved in syringyl lignin production in an angiosperm. The
 CC sequences are used to produce syringyl lignin in gymnosperms,
 CC especially the loblolly pine (Pinus taeda). It is necessary to
 CC remove much of the lignin from the fiber/lignin network of woody
 CC plants in paper manufacture. Greater proportions of syringyl lignin
 CC result in a higher delignification rate and hence a more efficient
 CC pulp mill operation.

XX Sequence 31 AA;

Query Match 76.5%; Score 39; DB 20; Length 31;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13
 III I I
 Db 7 CCGGIACIAC 16

RESULT 5

AAB19699
 ID AAB19699 standard; Peptide; 31 AA.

AC AAB19699;

DT 05-FEB-2001 (first entry)

XX 4-Coumarate CoA ligase conserved peptide.

DE Sweetgum; angiosperm; 4-coumarate CoA ligase; conifer;

KW loblolly pine; transgenic plant; lignin; paper; pulping.

XX Planta.

PN WO200058489-A2.

XX 05-OCT-2000.

PD 24-MAR-2000; 2000WO-US08083.

PF 26-MAR-1999; 99US-0277248.

PR (INTO) INT PAPER CO.

XX Chiang VL, Carraway DT;

DR WPI; 2000-647240/62.

XX Use of angiosperm conferyl aldehyde 5-hydroxylase which catalyzes
 PT 5-hydroxylation of conferyl aldehyde, for modifying lignin
 PT biosynthesis in gymnosperms, involves expressing the enzyme in a
 PT gymnosperm plant -

XX Example 1; Page 19; 123pp; English.

XX The present sequence is that of a conserved peptide of known
 CC 4-coumarate Co ligases (4CL). A primer based on this sequence
 CC was used in the identification of cDNA (see AAA88684) encoding
 CC sweetgum 4CL (see AAB19696), an enzyme involved in lignin
 CC biosynthetic pathways. An aim of the invention is to identify,
 CC sequence and clone specific genes from an angiosperm such as
 CC sweetgum that are involved in production of syringyl lignin, and
 CC to then introduce such genes into the genome of a gymnosperm, such
 CC as loblolly pine, to induce production of syringyl lignin and
 CC thereby provide enhanced pulpability to the wood structure.

XX Sequence 31 AA;

Query Match 76.5%; Score 39; DB 21; Length 31;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13
 III I I
 Db 7 CCGGIACIAC 16

RESULT 6

AAE05833
 ID AAE05833 standard; peptide; 31 AA.

XX AAE05833;

DT 24-SEP-2001 (first entry)

XX Liquidambar sp. angiosperm 4-coumarate CoA ligase conserved peptide #1.

KW Sweetgum; angiosperm; 4-coumarate CoA ligase; 4CL;
 KW syringyl lignin; gymnosperm; pulp; papermaking; guaiacyl lignin;
 KW delignification.

XX Liquidambar sp.

XX US6252135-B1.

PN 26-JUN-2001.

PD 16-DEC-1997; 97US-0991677.

PF 16-DEC-1996; 96US-0033381.

PR (INTO) INT PAPER CO.

DR WPI: 1998-179377/16.
DR N-PSDB: AAV21282.
XX Inhibitor of MyoD family proteins - useful for, e.g. treatment and
PT diagnosis of defects in myogenesis responsible for abnormal
PT development
XX
XX Disclosure: Pages 74-75; 92pp; English.
XX
XX The sequence is that of murine I-mfa (inhibitor of MyoD family)
CC protein. Probes from the gene sequence can be used for determining the
CC presence of an I-mf protein or analogue, or for detecting I-mf
CC agonist activity in a test substance. The sequence can be used
CC to treat, model and diagnose defects in myogenesis responsible
CC for abnormal development and disease conditions in mammals,
CC including humans. These include the human disease cleidocranial
CC dysplasia (CDD), rhabdomyosarcomas and other cancers affecting
CC muscle tissue in mammals.
XX
XX Sequence 246 AA;
SQ

Query Match 76.5%; Score 39; DB 19; Length 246;
Best Local Similarity 50.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CCGXXXCXXC 13
DB 203 CCGSGECADC 212

RESULT 9
ABB08605
ID ABB08605 standard; Protein: 349 AA.
XX
XX ABB08605;
AC
XX
DT 10-APR-2002 (first entry)
XX
XX Hirudin protein SEQ ID NO 15.
DE
XX
XX Fusion protein: hirudin; Hir: tick anticoagulant protein; TAP;
KW anticoagulant; blood; thrombin; factor Xa.
KW
XX
XX Hirudo medicinalis.
OS
XX
XX WO200204486-A2.
PN
XX
XX 17-JAN-2002.
PD
XX
XX 27-JUN-2001; 2001WO-EP07333.
PF
XX
XX 07-JUL-2000; 2000DE-1033195.
PR
XX
XX (AVET) AVENTIS PHARMA DEUT GMBH.
PA
XX
XX Habermann P;
PI
XX
XX WPI: 2002-154918/20.
DR
XX
XX New fusion protein of hirudin and tick anticoagulant protein, useful as
PT an anticoagulant -
PT
XX
XX Claim 3; Page 29-30; 36pp; German.
PS

The invention relates to a bifunctional fusion protein comprising
CC hirudin or its variant and tick anticoagulant protein (TAP) or its
CC variant. The fusion proteins have anticoagulant activity and are used to
CC inhibit coagulation of blood by acting to inhibit thrombin and inhibit
CC factor Xa. The present sequence is that of the hirudin protein, the first
CC 65 amino acids of which are given in figure 4.
CC
XX
XX Sequence 349 AA;
SQ

Query Match 76.5%; Score 39; DB 23; Length 349;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CCGXXXCXXC 13
DB 190 CCGAAACCGC 199

RESULT 10
AAG84970
ID AAG84970 standard; Protein: 919 AA.
XX
XX AAG84970;
AC
XX
DT 11-SEP-2001 (first entry)
XX
XX Shrimp white spot Bacilliform virus (WSBV) protein 61.
DE
XX
XX Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
KW antiviral agent; gene expression; antisense construct;
KW transgenic viral resistant shrimp.
KW
XX
XX White spot syndrome virus.
OS
XX
XX WO200138351-A2.
PN
XX
XX 31-MAY-2001.
PD
XX
XX 08-NOV-2000; 2000WO-US28888.
PF
XX
XX 24-NOV-1999; 99CN-0124717.
PR
XX
XX (PENY-) PE CORP NY.
PA
XX
XX (THIR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.
PA
XX
XX (SINO-) SINOGENOMAX CO LTD.
XX
XX Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;
PI
XX
XX WPI: 2001-355877/37.
DR
XX
XX N-PSDB; AAG62750.
DR

Primary nucleotide sequence of the shrimp white spot Bacilliform virus
PT (WSBV), useful for producing viral polypeptides that can be used to
PT screen for agents that are useful for treating WSBV infection -
XX
XX Claim 1; Figure 3; 626pp; English.
XX
XX The invention provides the primary nucleotide sequence of the WSBV genome
CC (AAG62689), predicted transcript sequences (AAG62689-AAG62839) and
CC encoded proteins (AAG84910-AAG85051) and oligonucleotide sequences
CC (AAG62840-63160) suitable for use as primers or probes. The nucleic acid
CC molecules and proteins of the invention are useful for diagnosis and
CC monitoring viral infection, in screens for antiviral agents and for
CC monitoring viral gene expression or activity during a treatment regimen.
CC The nucleic acid molecules are also useful as antisense constructs to
CC control viral gene expression in infected cells and tissues and to create
CC transgenic viral resistant shrimp.
XX
XX Sequence 919 AA;
SQ

Query Match 76.5%; Score 39; DB 22; Length 919;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CCGXXXCXXC 13
DB 412 CCGAKITSEC 421

RESULT 11
ABB59738
ID ABB59738 standard; Protein: 1231 AA.

```

XX ABB59738;
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 6006.
XX DE
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD
XX 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL03841.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 6006; 21pp + Sequence Listing; English.
XX PS
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511); expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins
XX CC (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1231 AA;
XX
XX Query Match 76.5%; Score 39; DB 22; Length 1231;
XX Best Local Similarity 50.0%; Pred. No. 5.3e+02;
XX Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX QY 4 CCGXXXCCXC 13
XX |||||
XX Db 232 CCGSFCDDC 241
XX
XX RESULT 12
XX AAE20581
XX ID AAE20581 standard; Protein; 1404 AA.
XX
XX AC AAE20581;
XX
XX DT 01-JUL-2002 (first entry)
XX
XX DE Proliferation potential protein (P2P).
XX
XX KW Proliferation potential protein; P2P; hnrnp; Rb1; cell proliferation;
XX KW tumour suppression; cancer; antisense gene therapy.
XX
XX OS Unidentified.
XX
XX FH Key Location/Qualifiers

```

```

FT Domain 61..101
FT /note= "Cysteine-rich domain"
FT Domain 79..97
FT /note= "Cell division sequence motif (CDSM)"
FT Region 718..725
FT /label= Nuclear_localisation_signal
FT Domain 753..908
FT /note= "Rb1 binding domain"
FT Region 795..811
FT /label= Nuclear_localisation_signal
FT Region 847..853
FT /label= Nuclear_localisation_signal
FT Domain 853..1404
FT /note= "hnRNP-associated domain"
FT Region 922..940
FT /label= Nuclear_localisation_signal
FT Region 1261..1277
FT /label= Nuclear_localisation_signal
XX US2002035080-A1.
XX
XX PD 21-MAR-2002.
XX
XX PF 16-MAR-2001; 2001US-0811045.
XX PR 27-SEP-1996; 96US-027568P.
XX PR 18-FEB-1997; 97US-0801308.
XX
XX PA (UYTE-) UNIV TENNESSEE RES CORP.
XX PI Scott RE;
XX
XX DR WPI; 2002-291590/33.
XX DR N-PSDB; AAD33333.
XX
XX PT New isolated proliferation potential protein nucleic acid and it's
XX PT antisense sequence, for repressing the proliferative potential of a
XX PT cell -
XX
XX PS Claim 6; Fig 2; 32pp; English.
XX
XX CC The present invention relates to proliferation potential proteins (P2P)
XX CC and polynucleotides encoding such proteins. P2P cDNAs encode proteins
XX CC with domains for hnRNP association and Rb1 binding. The interaction of
XX CC P2P cDNA products and Rb1 serve to modulate cell proliferation and/or
XX CC biological functions associated with tumour suppression by an RNA
XX CC processing mechanism. Antisense oligonucleotides to P2P polynucleotides
XX CC are used to repress the proliferative potential of a normal, abnormal
XX CC or cancer cell. Sequences of the invention are also used for antisense
XX CC gene therapy. The present sequence is P2P protein.
XX
XX SQ Sequence 1404 AA;
XX
XX Query Match 76.5%; Score 39; DB 23; Length 1404;
XX Best Local Similarity 50.0%; Pred. No. 5.7e+02;
XX Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
XX QY 4 CCGXXXCCXC 13
XX |||||
XX Db 76 CCGNSCDEC 85
XX
XX RESULT 13
XX AAU43831
XX ID AAU43831 standard; Protein; 80 AA.
XX
XX AC AAU43831;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE Propionibacterium acnes immunogenic protein #4727.
XX
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

```

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX WO2000181581-A2.
 PN
 XX
 XX
 XX 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 XX (CORI-) CORIXA CORP.
 XX
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59521.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 PT
 XX
 XX Example 1; SEQ ID NO 5026; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 80 AA;
 SQ
 Query Match 74.5%; Score 38; DB 22; Length 80;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 CCGXXXXXXC 13
 ||| |
 Db 17 CCGGACCCSC 26
 RESULT 14
 AAY70234
 ID AAY70234 standard; Protein; 153 AA.
 XX
 AC AAY70234;
 XX
 XX 06-JUN-2000 (first entry)
 DT
 DE Human RNA-associated protein-15 (RNAAP-15).
 XX
 XX RNA-associated protein; RNAAP; human; clone 1879135; cytostatic;

KW immunosuppressive; antiinflammatory; keratolytic; neuroprotective;
 KW antiarteriosclerotic; hepatotropic; antipsoriatic; virucide; anti-HIV;
 KW antiallergic; antirheumatic; antiarthritic; ophthalmological; autoimmune;
 KW antimicrobial; cell proliferative disorder; inflammation; cirrhosis;
 KW actinic keratosis; bursitis; arteriosclerosis; artherosclerosis;
 KW hepatitis; myelofibrosis; primary thrombocythemia; psoriasis; cancer;
 KW mixed connective tissue disease; MCTD; HIV; uveitis; Crohn's disease;
 KW allergy; rheumatoid arthritis; parasitic infection.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 5
 FT /note= "Potential phosphorylation site"
 FT Domain 26..43
 FT /label= Zinc_finger
 FT Modified-site 43
 FT /note= "Potential phosphorylation site"
 FT Modified-site 59
 FT /note= "Potential phosphorylation site"
 FT Modified-site 82
 FT /note= "Potential phosphorylation site"
 FT Modified-site 112
 FT /note= "Potential phosphorylation site"
 FT Modified-site 113
 FT /note= "Potential phosphorylation site"
 FT Modified-site 114
 FT /note= "Potential phosphorylation site"
 FT Modified-site 145
 FT /note= "Potential phosphorylation site"
 FT
 XX WO200011171-A2.
 PN
 XX 02-MAR-2000.
 XX
 XX 20-AUG-1999; 99WO-US19361.
 XX
 PR 21-AUG-1998; 98US-0097550.
 PR 12-JAN-1999; 99US-0115639.
 XX
 XX (INCY-) INCYTE PHARM INC.
 XX
 XX Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
 PI Patterson C, Baughn MR, Lal P, Bandman O, Reddy R, Azimzai Y;
 PI Shih LL, Yang J, Lu DAM;
 XX
 XX WPI; 2000-237651/20.
 DR N-PSDB; AAZ51264.
 XX
 XX Human RNA-associated proteins useful in diagnosing, treating and
 PT preventing cell proliferative, autoimmune, inflammatory and infectious
 PT disorders.
 XX
 XX Claim 1; Page 94-95; 123pp; English.
 PS
 XX The present amino acid sequence is the human RNA-associated protein-15
 CC (RNAAP-15), identified in incyte clone 1879135, derived from LEUKNOT03
 CC library. It is expressed in reproductive, dermatologic, gastrointestinal,
 CC musculoskeletal and haematopoietic/immune tissues. It has cytostatic,
 CC immunosuppressive, antiinflammatory, antiarteriosclerotic, hepatotropic,
 CC keratolytic, neuroprotective, antipsoriatic, anti-HIV, antiallergic,
 CC antirheumatic, virucide, antiarthritic, ophthalmological and antimicrobial
 CC activity. RNAAP antibodies are useful for diagnosis of diseases
 CC associated with altered expression or activity of RNAAP. It is used to
 CC treat cell proliferative, autoimmune, inflammatory and infectious
 CC disorders, like actinic keratosis, bursitis, arteriosclerosis,
 CC artherosclerosis, cirrhosis, hepatitis, myelofibrosis, mixed connective
 CC tissue disease (MCTD), psoriasis, primary thrombocythemia and cancer,
 CC HIV, allergies, rheumatoid arthritis, uveitis, Crohn's disease, and
 CC bacterial, viral and parasitic infections.
 XX
 XX Sequence 153 AA;

Query Match 74.5%; Score 38; DB 21; Length 153;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 CCGXXCXXC 13
||| |
Db 141 CCGNSYCDEC 150

RESULT 15

ABB69415
ID ABB69415 standard; Protein; 237 AA.

XX
AC ABB69415;

XX
DT 26-MAR-2002 (first entry)

XX
DE Drosophila melanogaster polypeptide SEQ ID NO 35037.

XX
KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

XX
OS Drosophila melanogaster.

XX
PN WO200171042-A2.

XX
PD 27-SEP-2001.

XX
PF 23-MAR-2001; 2001WO-US09231.

XX
PR 23-MAR-2000; 2000US-191637P.

XX
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.

XX
PI Venter JC, Adams M, Li PWD, Myers EW;

XX
DR WPI: 2001-656860/75.

XX
DR N-PSDB; ABL13518.

XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX
PS Disclosure; SEQ ID NO 35037; 21pp + Sequence Listing; English.

XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLL16176-ABLL30511), expressed DNA
CC sequences (ABLL1840-ABLL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 237 AA;

Query Match 74.5%; Score 38; DB 22; Length 237;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 CCGXXCXXC 13
||| |
Db 165 CCGKQKCTGC 174

Search completed: January 29, 2003, 09:30:57
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:57 ; Search time 5.56818 Seconds
(without alignments)
73.978 Million cell updates/sec

Title: US-09-580-201A-1

Perfect score: 51

Sequence: 1 XXXCGXXXXXXC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	40	78.4	460	1	US-08-689-974-5
2	40	78.4	460	2	US-09-058-376-5
3	40	78.4	1345	3	US-08-977-767-3
4	40	78.4	1917	4	US-09-627-650B-5
5	40	78.4	1917	4	US-09-436-063C-5
6	39	76.5	57	1	US-07-609-716-56
7	39	76.5	246	2	US-07-609-931-2
8	39	76.5	1404	4	US-08-801-308-1
9	39	76.5	2508	4	US-09-627-650B-7
10	39	76.5	2508	4	US-09-436-063C-7
11	39	76.5	2544	4	US-09-627-650B-3
12	39	76.5	2544	4	US-09-436-063C-3
13	39	76.5	2601	4	US-09-627-650B-9
14	39	76.5	2601	4	US-09-436-063C-9
15	38	74.5	38	1	US-07-609-716-53
16	38	74.5	801	1	US-07-906-349A-6
17	37	72.5	41	5	PCT-US96-01720-7
18	37	72.5	55	3	US-08-476-509B-46
19	37	72.5	1417	4	US-08-900-230-3
20	36	70.6	21	2	US-08-754-431A-10
21	36	70.6	2211	4	US-09-738-884-1
22	35	68.6	1400	4	US-08-630-915A-37
23	34	66.7	2088	4	US-09-548-372D-13
24	34	66.7	2088	4	US-09-548-367D-13
25	33	64.7	22	1	US-08-423-398B-27
26	33	64.7	50	4	US-08-900-230-8
27	33	64.7	143	4	US-08-990-823-112

28	33	64.7	1128	4	US-09-627-650B-11	Sequence 11, Appl
29	33	64.7	1128	4	US-09-436-063C-11	Sequence 11, Appl
30	33	64.7	1652	4	US-09-627-650B-1	Sequence 1, Appl
31	33	64.7	1652	4	US-09-436-063C-1	Sequence 1, Appl
32	32	62.7	27	4	US-09-073-407-1	Sequence 1, Appl
33	32	62.7	27	4	US-09-073-407-2	Sequence 2, Appl
34	32	62.7	27	4	US-09-073-407-3	Sequence 3, Appl
35	32	62.7	27	4	US-09-073-407-4	Sequence 4, Appl
36	32	62.7	27	4	US-09-073-407-5	Sequence 5, Appl
37	32	62.7	27	4	US-09-073-407-6	Sequence 6, Appl
38	32	62.7	27	4	US-09-073-407-7	Sequence 7, Appl
39	32	62.7	27	4	US-09-073-407-8	Sequence 8, Appl
40	32	62.7	43	6	5212296-3	Patent No. 5212296
41	32	62.7	47	1	US-08-451-947-94	Sequence 94, Appl
42	32	62.7	47	1	US-08-424-826A-94	Sequence 94, Appl
43	32	62.7	47	3	US-08-482-085B-91	Sequence 91, Appl
44	32	62.7	47	3	US-08-928-694-94	Sequence 94, Appl
45	32	62.7	47	5	PCT-US91-06950-94	Sequence 94, Appl

ALIGNMENTS

RESULT 1
US-08-689-974-5
; Sequence 5, Application US/08689974
; Patent No. 5776732
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Murray, Lynn E.
; TITLE OF INVENTION: NOVEL HUMAN INDUCED TUMOR PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,974
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0113 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-853-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 310100
US-08-689-974-5

Query Match 78.4%; Score 40; DB 1; Length 460;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXXXC 13
||| | |

Db 13 CCGSAGCALC 22

RESULT 2

US-09-058-376-5
; Sequence 5, Application US/09058376
; Patent No. 6080841
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Murray, Lynn E.
; TITLE OF INVENTION: NOVEL HUMAN INDUCED TUMOR PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,376
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,974
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0113 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 310100

Query Match 78.4%; Score 40; DB 3; Length 460;

Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13

Db 13 CCGSAGCALC 22

RESULT 3

US-08-977-767-3
; Sequence 3, Application US/08977767
; Patent No. 5972684
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Greenwald, Sara
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive

; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,767
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0423 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1532042

US-08-977-767-3

Query Match 78.4%; Score 40; DB 2; Length 1345;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13

Db 1089 CCGTAACGGC 1098

RESULT 4

US-09-627-650B-5
; Sequence 5, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1917
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

US-09-627-650B-5

Query Match 78.4%; Score 40; DB 4; Length 1917;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13

```
Db 17 CCGTTACCTC 26
      ||| | |
Query Match 78.4%; Score 40; DB 4; Length 1917;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
      ||| | |
Db 17 CCGTTACCTC 26
      ||| | |

RESULT 6
US-07-609-716-56
; Sequence 56, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Query Match 76.5%; Score 39; DB 2; Length 246;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
      ||| | |
Db 203 CCGGECADC 212
      ||| | |

RESULT 8
US-08-801-308-1
; Sequence 1, Application US/08801308
; Patent No. 6368790
; GENERAL INFORMATION:
; APPLICANT: Scott, Robert E.
; TITLE OF INVENTION: CDNA ENCODING P2P PROTEINS AND USE OF
; TITLE OF INVENTION: P2P CDNA-DERIVED ANTIBODIES AND ANTISENSE REAGENTS IN
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-07-609-716-56

Query Match 76.5%; Score 39; DB 1; Length 57;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
      ||| | |
Db 28 CCGGGCTAC 37
      ||| | |

RESULT 7
US-08-704-931-2
; Sequence 2, Application US/08704931
; Patent No. 5885797
; GENERAL INFORMATION:
; APPLICANT: Chen, Chao-Min (Amy)
; APPLICANT: Kraut, No. 5885797bert
; APPLICANT: Groudiue, Mark
; APPLICANT: Weintraub, Harold
; TITLE OF INVENTION: No. 5885797el DNA Sequences Encoding Proteins
; TITLE OF INVENTION: Involved in Myogenesis
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stratton Balliew, PLLC
; STREET: 1218 Third Avenue, Suite 1313
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,931
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: King, Jeffrey J
; REGISTRATION NUMBER: 38,515
; REFERENCE/DOCKET NUMBER: HUI1.P02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-683-1496
; TELEFAX: 206-682-0446
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-704-931-2

Query Match 76.5%; Score 39; DB 2; Length 246;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
      ||| | |
Db 203 CCGGECADC 212
      ||| | |

RESULT 8
US-08-801-308-1
; Sequence 1, Application US/08801308
; Patent No. 6368790
; GENERAL INFORMATION:
; APPLICANT: Scott, Robert E.
; TITLE OF INVENTION: CDNA ENCODING P2P PROTEINS AND USE OF
; TITLE OF INVENTION: P2P CDNA-DERIVED ANTIBODIES AND ANTISENSE REAGENTS IN
```

;; TITLE OF INVENTION: DETERMINING THE PROLIFERATIVE POTENTIAL OF NORMAL,
;; TITLE OF INVENTION: ABNORMAL AND CANCER CELLS IN ANIMALS AND HUMANS
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Weiser & Associates, P.C.
;; STREET: 230 S. Fifteenth Street, Suite 500
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19102

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/801,308

;; FILING DATE: 18-FEB-1997

;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Weiser, Gerard J.

;; REGISTRATION NUMBER: 19,763

;; REFERENCE/DOCKET NUMBER: 372.6435P

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 215-875-8383

;; TELEFAX: 215-875-8394

;; INFORMATION FOR SEQ ID NO: 1:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 1404 amino acids

;; TYPE: amino acid

;; TOPOLOGY: linear

;; MOLECULE TYPE: protein

US-08-801-308-1

Query Match 76.5%; Score 39; DB 4; Length 1404;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13

Db 76 CCGNSSCDEC 85

RESULT 9

US-09-627-650B-7

; Sequence 7, Application US/09627650B

; Patent No. 6406872

; GENERAL INFORMATION:

; APPLICANT: Bamber, Bruce

; APPLICANT: Jorgensen, Erik

; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and

; TITLE OF INVENTION: Methods Related Thereto

; FILE REFERENCE: 21101.000903

; CURRENT APPLICATION NUMBER: US/09/627,650B

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 09/436,063

; PRIOR FILING DATE: 1999-11-08

; PRIOR APPLICATION NUMBER: 60/107,727

; PRIOR FILING DATE: 1998-11-09

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 2508

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-09-627-650B-7

Query Match 76.5%; Score 39; DB 4; Length 2508;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13

Db 76 CCGNSSCDEC 85

Db 830 CCGAGACCTC 839

RESULT 10

US-09-436-063C-7

; Sequence 7, Application US/09436063C

; Patent No. 6407210

; GENERAL INFORMATION:

; APPLICANT: Bamber, Bruce

; APPLICANT: Jorgensen, Erik

; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and

; TITLE OF INVENTION: Methods Related Thereto

; FILE REFERENCE: P-1095Corrected

; CURRENT APPLICATION NUMBER: US/09/436,063C

; CURRENT FILING DATE: 1999-11-08

; PRIOR APPLICATION NUMBER: 60/107727

; PRIOR FILING DATE: 1998-11-09

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

; LENGTH: 2508

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-09-436-063C-7

Query Match 76.5%; Score 39; DB 4; Length 2508;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13

Db 830 CCGAGACCTC 839

RESULT 11

US-09-627-650B-3

; Sequence 3, Application US/09627650B

; Patent No. 6406872

; GENERAL INFORMATION:

; APPLICANT: Bamber, Bruce

; APPLICANT: Jorgensen, Erik

; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and

; TITLE OF INVENTION: Methods Related Thereto

; FILE REFERENCE: 21101.000903

; CURRENT APPLICATION NUMBER: US/09/627,650B

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 09/436,063

; PRIOR FILING DATE: 1999-11-08

; PRIOR APPLICATION NUMBER: 60/107,727

; PRIOR FILING DATE: 1998-11-09

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 2544

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-09-627-650B-3

Query Match 76.5%; Score 39; DB 4; Length 2544;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13

Db 830 CCGAGACCTC 839

RESULT 12

US-09-436-063C-3

; Sequence 3, Application US/09436063C

; Patent No. 6407210

; GENERAL INFORMATION:

; APPLICANT: Bamber, Bruce

; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2544
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-3

Query Match 76.5%; Score 39; DB 4; Length 2544;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
||| | |
Db 830 CCGAGACCTC 839

RESULT 13

; Sequence 9, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-9

Query Match 76.5%; Score 39; DB 4; Length 2601;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
||| | |
Db 830 CCGAGACCTC 839

RESULT 14

; Sequence 9, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-9

Query Match 76.5%; Score 39; DB 4; Length 2601;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
||| | |
Db 830 CCGAGACCTC 839

RESULT 15

; Sequence 53, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-07-609-716-53

Query Match 74.5%; Score 38; DB 1; Length 38;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
||| | |
Db 1 CCGGTCCTC 10

Search completed: January 29, 2003, 09:33:40
Job time : 8.56818 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 10.9773 Seconds
(without alignments)
262.785 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCCGXXXXXCX 14
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	40	78.4	460	11 Q63175	Q63175 rattus norv
2	39	76.5	170	6 Q9T27	Q9T27 bos taurus
3	39	76.5	170	6 Q9T26	Q9T26 oviss aries
4	39	76.5	188	2 Q93F29	Q93F29 methylophag
5	39	76.5	246	4 Q99750	Q99750 homo sapien
6	39	76.5	246	11 Q99750	Q99750 mus musculu
7	39	76.5	286	3 Q06651	Q06651 saccharomyc
8	39	76.5	350	5 Q988U4	Q988U4 trypanosoma
9	39	76.5	721	12 Q91L5	Q91L5 white spot
10	39	76.5	1009	5 Q8VAC3	Q8VAC3 white spot
11	39	76.5	1190	5 Q62235	Q62235 caenorhabdi
12	39	76.5	1231	5 Q9XZ21	Q9XZ21 drosophila
13	38	74.5	237	5 Q9V913	Q9V913 drosophila
14	38	74.5	258	12 Q36453	Q36453 lymantria d
15	38	74.5	326	2 Q9WX86	Q9WX86 clostridium
16	38	74.5	326	16 Q8XJF9	Q8XJF9 clostridium

17	38	74.5	326	16 Q8RA96	Q8RA96 thermoanaer
18	38	74.5	328	16 Q97H26	Q97H26 clostridium
19	38	74.5	363	5 Q9N5Y3	Q9N5Y3 caenorhabdi
20	38	74.5	948	4 Q15290	Q15290 homo sapien
21	38	74.5	1023	5 Q9XYDA	Q9XYDA dictyosteli
22	38	74.5	1591	11 F97868	F97868 mus musculu
23	38	74.5	1616	4 Q96PH3	Q96PH3 homo sapien
24	37	72.5	151	12 Q10614	Q10614 heliocoverpa
25	37	72.5	167	11 Q9D122	Q9D122 mus musculu
26	37	72.5	168	11 Q9D732	Q9D732 mus musculu
27	37	72.5	186	11 Q64526	Q64526 mus musculu
28	37	72.5	191	11 Q9D3H7	Q9D3H7 mus musculu
29	37	72.5	195	11 Q9D141	Q9D141 mus musculu
30	37	72.5	234	12 Q8QL66	Q8QL66 mamestra co
31	37	72.5	244	12 Q9IBQ9	Q9IBQ9 spodoptera
32	37	72.5	263	16 Q8XKG1	Q8XKG1 clostridium
33	37	72.5	264	16 Q97IX6	Q97IX6 clostridium
34	37	72.5	285	12 Q99H38	Q99H38 heliocoverp
35	37	72.5	285	12 Q8V5X4	Q8V5X4 heliocoverpa
36	37	72.5	289	12 Q91BK7	Q91BK7 spodoptera
37	37	72.5	312	5 Q45799	Q45799 caenorhabdi
38	37	72.5	446	4 Q96BQ3	Q96BQ3 homo sapien
39	37	72.5	550	10 Q23187	Q23187 arabidopsis
40	37	72.5	693	6 Q27960	Q27960 bos taurus
41	37	72.5	1968	3 Q8X0C5	Q8X0C5 neurospora
42	36	70.6	64	5 Q9BPE9	Q9BPE9 conus penna
43	36	70.6	102	11 Q9D702	Q9D702 mus musculu
44	36	70.6	113	11 Q9D228	Q9D228 mus musculu
45	36	70.6	122	11 Q9D230	Q9D230 mus musculu

ALIGNMENTS

RESULT 1

Q63175 PRELIMINARY; PRT; 460 AA.
ID Q63175;
AC Q63175;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Developmentally regulated protein.
GN TPOL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=TELENCEPHALON;
RX MEDLINE=97465542; PubMed=9326262;
RA Krueger W.H., Gonye G.E., Madison D.L., Murray K.E., Kumar M.,
RA Spoerel N., Pfeiffer S.E.;
RT "TPOL, a member of a novel protein family, is developmentally
regulated in cultured oligodendrocytes.";
RL J. Neurochem. 69:1343-1355(1997).
CC -!- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
DR EMBL; L20319; AAA41097.1;
DR InterPro; IPR000564; 2Fe2S-ferredoxin.
DR InterPro; IPR005016; TMS_TDE.
DR Pfam; PF03348; TMS_TDE; 1.
DR PROSITE; PS00197; 2FE2S-FERREDOXIN; 1.
KW Iron-sulfur.
SQ SEQUENCE 460 AA; 51855 MW; CB9DA60F66282FB3 CRC64;

Query Match 78.4%; Score 40; DB 11; Length 460;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXXCX 13

DB 13 CCGSAGCALC 22

Query Match 76.5%; Score 39; DB 6; Length 170;

```
DR InterPro; IPR001368; TNFR_c6.
DR InterPro; IPR001007; VWF_C.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01208; VWF; UNKNOWN_1.
SQ SEQUENCE 246 AA; 25029 MW; 9330E52141B79E5A CRC64;

Query Match 76.5%; Score 39; DB 4; Length 246;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXC 13
Db 203 CCGSGECADC 212

RESULT 6
O99JM9
ID Q99JM9 PRELIMINARY; PRT; 246 AA.
AC Q99JM9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MyD family inhibitor.
GN MFI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006018; AAH06018.1; -.
DR EMBL; BC010259; AAH10259.1; -.
DR MGD; MGI-107687; Mfi.
DR InterPro; IPR001368; TNFR_c6.
DR InterPro; IPR001007; VWF_C.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01208; VWF; UNKNOWN_1.
SQ SEQUENCE 246 AA; 25270 MW; 138A37A2B4A10463 CRC64;

Query Match 76.5%; Score 39; DB 11; Length 246;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXC 13
Db 203 CCGSGECADC 212

RESULT 7
Q06651
ID Q06651 PRELIMINARY; PRT; 286 AA.
AC Q06651;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to proteins involved in vacuolar function: S. CEREVISIAE
DE VACIP.
GN PIB1 OR D9740.20 OR YDR313C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
```

```
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Taich A., Trevaskis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ding H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Jia Y., Cherry J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28374; AAB64749.1; -.
DR SGD; S0002721; PIB1.
DR InterPro; IPR000306; Znf_FYVE.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00084; FYVE; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 286 AA; 32675 MW; E7E9DCDEC9BAA183 CRC64;

Query Match 76.5%; Score 39; DB 3; Length 286;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXC 13
Db 41 CCGRIFCSCC 50

RESULT 8
Q9N8U4
ID Q9N8U4 PRELIMINARY; PRT; 350 AA.
AC Q9N8U4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical 38.0 kDa protein.
GN CHL143.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TREU927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359782; CAB95439.1; -.
KW Hypothetical protein.
SQ SEQUENCE 350 AA; 37995 MW; 98796C546331D2A4 CRC64;

Query Match 76.5%; Score 39; DB 5; Length 350;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXC 13
Db 234 CCGAIMCGAC 243

RESULT 9
Q91LL5
ID Q91LL5 PRELIMINARY; PRT; 721 AA.
AC Q91LL5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE ORF24.
 OS White spot syndrome virus (WSSV).
 OC Viruses; unclassified viruses.
 OX NCBI_TaxID=92652;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21342572; PubMed=11448154;
 RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
 RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
 RT "The white spot syndrome virus DNA genome sequence.";
 RL Virology 286:7-22(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
 RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF369029; AAK77693.1; -.
 SQ SEQUENCE 721 AA; 81137 MW; B1F9850F3CEA494D CRC64;

Query Match 76.5%; Score 39; DB 12; Length 721;
 Best Local Similarity 50.0%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
 ||| |
 Db 212 CCGAKICSEC 221

RESULT 10
 Q8VAC3 PRELIMINARY; PRT; 1009 AA.
 AC Q8VAC3;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE wsv502.
 OS White spot syndrome virus (WSSV).
 OC Viruses; unclassified viruses.
 OX NCBI_TaxID=92652;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21548311; PubMed=11689662;
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
 RT "Complete genome sequence of the shrimp white spot bacilliform
 RT virus";
 RL J. Virol. 75:11811-11820(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF332093; AAL33503.1; -.
 SQ SEQUENCE 1009 AA; 113479 MW; 5E9BC87638066D6 CRC64;

Query Match 76.5%; Score 39; DB 12; Length 1009;
 Best Local Similarity 50.0%; Pred. No. 63;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
 ||| |
 Db 500 CCGAKICSEC 509

RESULT 11
 O62235 PRELIMINARY; PRT; 1190 AA.
 AC O62235;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE F36F2.3 protein.
 GN F36F2.3.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Cottage A.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RT none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 DR EMBL; Z81532; CAB04326.2; -.
 DR InterPro; IPR001357; BRCT.
 DR InterPro; IPR004274; NIF.
 DR InterPro; IPR001878; Znf_CCHC.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00533; BRCT; 1.
 DR Pfam; PF03031; NIF; 1.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00343; Znf_C2HC; 1.
 SQ SEQUENCE 1190 AA; 137239 MW; B219D269962432AB CRC64;

Query Match 76.5%; Score 39; DB 5; Length 1190;
 Best Local Similarity 50.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
 ||| |
 Db 323 CCGNSYCADC 332

RESULT 12
 O9XZ21 PRELIMINARY; PRT; 1231 AA.
 AC O9XZ21;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE BCDA; LD21643 protein.
 GN BCDA; LD21643 OR CG3231.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX STRAIN=BERKELEY;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., B.D.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kelush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN
RP
SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcania T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm E.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AE003463; AAF47162.1; -;
DR EMBL: AFI32117; AAD34765.1; -;
DR FlyBase; FBgn0027522; BcDNA:LD21643.
DR InterPro; IPR001878; Znf_CCHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR Pfam; PF00098; zf-CCHC; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00343; Znf_C2HC; 1.
DR KW Hypothetical protein; Zinc-finger.
SQ SEQUENCE 1231 AA; 139038 MW; 4A991FF172BDE4B0 CRC64;

Query Match 76.5%; Score 39; DB 5; Length 1231;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXCC 13
||| | |
Db 232 CCGSFCDDC 241

RESULT 13
Q9V9I3
ID Q9V9I3 PRELIMINARY; PRT; 237 AA.
AC Q9V9I3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG14468 protein.
GN TS42A OR CG14468.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003785; AAF57306.1; -;
DR FlyBase; FBgn0033042; Tsp42A.
DR InterPro; IPR000301; Transmem_4.
SQ SEQUENCE 237 AA; 27034 MW; 43E964F0E721D962 CRC64;

Query Match 74.5%; Score 38; DB 5; Length 237;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCCXCC 13
||| | |
Db 165 CCGKQNCCTGC 174

RESULT 14
O36453
ID O36453 PRELIMINARY; PRT; 258 AA.
AC O36453; O9YMW3;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Immediate early 0 protein (Immediate early transactivator 0).
GN IE-0.
OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445058; PubMed=9300047;
RA Pearson M.N., Rohmann G.F.;
RT "Splicing is required for transactivation by the immediate early gene
RT 1 of the Lymantria dispar multinucleocapsid nuclear polyhedrosis
RT virus";
RL Virology 235:153-165(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99124785; PubMed=9887315;
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,

RA Slavicek J.M., Rohrmann G.F.;
 RT "Sequence and analysis of the genome of a baculovirus pathogenic for
 RT Lymantria dispar";
 RL Virology 253:17-34(1999).
 RN [3]

RP SEQUENCE FROM N.A.
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J., Rohrmann G.F.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF006656; AAC58234.1; -;
 DR EMBL: AF081810; AAC70206.1; -;
 DR InterPro: IPR001841; Znf_ring.
 DR SMART: SM00184; RING; 1.
 SQ SEQUENCE 258 AA; 29395 MW; AC564CDF92282BAD CRC64;

Query Match 74.5%; Score 38; DB 12; Length 258;
 Best Local Similarity 50.0%; Pred. No. 44;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
 Db 209 CCGYRVCNAC 218

RESULT 15

Q9WX86
 ID Q9WX86 PRELIMINARY; PRT; 326 AA.
 AC Q9WX86;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Electron transfer subunit protein.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_taxid=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PB6K;
 RX MEDLINE=20090462; PubMed=10627036;
 RA Fujinaga K., Taniguchi Y., Sun Y., Katayama S., Minami J.,
 RA Matsushita O., Okabe A.;
 RT "Analysis of genes involved in nitrate reduction in Clostridium
 RT perfringens";
 RL Microbiology 145:3377-3387(1999).
 CC -!- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
 DR EMBL: AB017192; BAA74788.1; -;
 DR InterPro: IPR000564; 2Fe2S_ferredoxin.
 DR InterPro: IPR001281; Rieske.
 DR PROSITE: PS00197; 2FE2S_FERREDOXIN; 1.
 DR PROSITE: PS00200; RIESKE_2; UNKNOWN_1.
 KW Iron-sulfur.
 SQ SEQUENCE 326 AA; 36522 MW; 2BDC1798F376FD65 CRC64;

Query Match 74.5%; Score 38; DB 2; Length 326;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
 Db 288 CCGGVCGSC 297

Search completed: January 29, 2003, 09:32:14
 Job time : 12.9773 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:29:32 ; Search time 3.34091 Seconds
(without alignments)
84.558 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCCGXXXXXXCX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	82.4	1422	10	US-09-735-933-1
2	41	80.4	1435	9	US-10-174-590-581
3	41	80.4	1435	9	US-10-176-758-581
4	41	80.4	1435	9	US-10-175-737-581
5	41	80.4	1435	12	US-10-052-586-581
6	41	80.4	1497	10	US-09-060-854B-2
7	40	78.4	24	8	US-08-775-765-8
8	40	78.4	210	9	US-10-114-893-107
9	40	78.4	3907	9	US-10-029-217A-24
10	40	78.4	4679	10	US-09-804-898-2
11	39	76.5	41	8	US-08-969-137-1
12	39	76.5	41	8	US-08-969-137-4
13	39	76.5	169	10	US-09-925-301-950
14	39	76.5	1404	10	US-09-811-045A-1
15	38	74.5	914	9	US-09-975-143-47
16	38	74.5	1184	12	US-10-153-064-89
17	38	74.5	1200	10	US-09-826-508-3
18	38	74.5	1300	9	US-10-174-590-269
19	38	74.5	1300	9	US-10-176-758-269

20	38	74.5	1300	9	US-10-175-737-269
21	38	74.5	1300	12	US-10-052-586-269
22	37	72.5	720	10	US-09-756-071B-20
23	36	70.6	975	10	US-09-886-055-431
24	36	70.6	2211	9	US-10-096-981-1
25	35	68.6	1400	10	US-09-879-957-37
26	35	68.6	1743	9	US-10-174-590-451
27	35	68.6	1743	9	US-10-176-758-451
28	35	68.6	1743	9	US-10-175-737-451
29	35	68.6	1743	12	US-10-052-586-451
30	33	64.7	143	9	US-09-996-634-112
31	33	64.7	228	9	US-09-791-932-17
32	33	64.7	480	10	US-09-823-936-16
33	33	64.7	4440	9	US-10-174-590-525
34	33	64.7	4440	9	US-10-176-758-525
35	33	64.7	4440	9	US-10-175-737-525
36	33	64.7	4440	12	US-10-052-586-525
37	32	62.7	23	10	US-09-860-793-9
38	32	62.7	38	9	US-09-834-988A-5
39	32	62.7	47	8	US-08-450-842-94
40	32	62.7	74	10	US-09-731-242A-4
41	32	62.7	90	10	US-09-860-793-5
42	32	62.7	134	10	US-09-864-761-39564
43	32	62.7	544	9	US-09-791-932-40
44	32	62.7	1076	9	US-10-028-072-219
45	32	62.7	1253	10	US-09-901-106-3

ALIGNMENTS

RESULT 1

US-09-735-933-1
; Sequence 1, Application US/09735933
; Patent No. US20020052034A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; OF INVENTION: THEREOF
; FILE REFERENCE: CL000863
; CURRENT APPLICATION NUMBER: US/09/735,933
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1422
; TYPE: PRT
; ORGANISM: Human
US-09-735-933-1

Query Match 82.4%; Score 42; DB 10; Length 1422;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXXXXC 13
||| |
Db 275 CCGTATCAAC 284

RESULT 2

US-10-174-590-581
; Sequence 581, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria

```
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC42
; CURRENT APPLICATION NUMBER: US/10/174,590
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 581
; LENGTH: 1435
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-581

Query Match      80.4%; Score 41; DB 9; Length 1435;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      4 CCGXXXCXXC 13
      ||| | |
Db      899 CCGTTTCGAC 908

RESULT 3
US-10-176-758-581
; Sequence 581, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 581
; LENGTH: 1435
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-581

Query Match      80.4%; Score 41; DB 9; Length 1435;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      4 CCGXXXCXXC 13
      ||| | |
Db      899 CCGTTTCGAC 908

RESULT 4
US-10-175-737-581
; Sequence 581, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
```

```
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 581
; LENGTH: 1435
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-581

Query Match      80.4%; Score 41; DB 9; Length 1435;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy      4 CCGXXXCXXC 13
      ||| | |
Db      899 CCGTTTCGAC 908

RESULT 5
US-10-052-586-581
; Sequence 581, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
```

;	PRIOR FILING DATE:	1998-05-07
;	PRIOR APPLICATION NUMBER:	60/084643
;	PRIOR FILING DATE:	1998-05-07
;	PRIOR APPLICATION NUMBER:	60/085573
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/085579
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/085580
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/085582
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/085700
;	PRIOR FILING DATE:	1998-05-15
;	PRIOR APPLICATION NUMBER:	60/086023
;	PRIOR FILING DATE:	1998-05-18
;	PRIOR APPLICATION NUMBER:	60/086392
;	PRIOR FILING DATE:	1998-05-22
;	PRIOR APPLICATION NUMBER:	60/086486
;	PRIOR FILING DATE:	1998-05-22
;	PRIOR APPLICATION NUMBER:	60/087098
;	PRIOR FILING DATE:	1998-05-28
;	PRIOR APPLICATION NUMBER:	60/087208
;	PRIOR FILING DATE:	1998-05-28
;	PRIOR APPLICATION NUMBER:	60/087609
;	PRIOR FILING DATE:	1998-06-02
;	PRIOR APPLICATION NUMBER:	60/087759
;	PRIOR FILING DATE:	1998-06-02
;	PRIOR APPLICATION NUMBER:	60/087827
;	PRIOR FILING DATE:	1998-06-03
;	PRIOR APPLICATION NUMBER:	60/088025
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088028
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088029
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088033
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088167
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088202
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088212
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088217
;	PRIOR FILING DATE:	1998-06-05
;	PRIOR APPLICATION NUMBER:	60/088326
;	PRIOR FILING DATE:	1998-06-04
;	PRIOR APPLICATION NUMBER:	60/088655
;	PRIOR FILING DATE:	1998-06-09
;	PRIOR APPLICATION NUMBER:	60/088722
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088738
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088740
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088811
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088824
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088825
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088826
;	PRIOR FILING DATE:	1998-06-10
;	PRIOR APPLICATION NUMBER:	60/088861
;	PRIOR FILING DATE:	1998-06-11
;	PRIOR APPLICATION NUMBER:	60/088863
;	PRIOR FILING DATE:	1998-06-11
;	PRIOR APPLICATION NUMBER:	60/088876
;	PRIOR FILING DATE:	1998-06-11
;	PRIOR APPLICATION NUMBER:	60/089090
;	PRIOR FILING DATE:	1998-06-12
;	PRIOR APPLICATION NUMBER:	60/089105
;	PRIOR FILING DATE:	1998-06-12

```
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

Query Match      80.4%; Score 41; DB 12; Length 1435;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY  4 CCGXXCXXC 13
   ||| | |
Db  899 CCGTTTCGAC 908

RESULT 6
US-09-060-854B-2
; Sequence 2, Application US/09060854B
; Patent No. US20020081703A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David Aaron
; TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical
; TITLE OF INVENTION: Applications and for Reducing the Allergenicity of No. US2002008
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: GC532
; CURRENT APPLICATION NUMBER: US/09/060,854B
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: B. amyloliquefaciens
US-09-060-854B-2

Query Match      80.4%; Score 41; DB 10; Length 1497;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY  4 CCGXXCXXC 13
   ||| | |
Db  841 CCGTTGCATC 850

RESULT 7
US-08-775-765-8
; Sequence 8, Application US/08775765C
; Patent No. US20010010821A1
; GENERAL INFORMATION:
; APPLICANT: Kelleher, Dermot
; APPLICANT: Windle, Henry
; APPLICANT: Byrne, William
; APPLICANT: McManus, Ross
; TITLE OF INVENTION: Helicobacter Proteins and Vaccines
; FILE REFERENCE: 08/775 765
; CURRENT APPLICATION NUMBER: US/08/775,765C
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IE 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IE 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Mus musculus
```

```
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-775-765-8

Query Match      78.4%; Score 40; DB 8; Length 24;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY  4 CCGXXCXXC 13
   ||| | |
Db  8 CCGTYTCYTC 17

RESULT 8
US-10-114-893-107
; Sequence 107, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R. A.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-107

Query Match      78.4%; Score 40; DB 9; Length 210;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY  4 CCGXXCXXC 13
   ||| | |
Db  13 CCGSAGCSLC 22

RESULT 9
US-10-029-217A-24
; Sequence 24, Application US/10029217A
; Patent No. US20020164735A1
; GENERAL INFORMATION:
; APPLICANT: OLSON, ERIC N.
; APPLICANT: WANG, DA-ZHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
; TITLE OF INVENTION: NUCLEAR REGULATORY FACTOR
; FILE REFERENCE: U7SD:695US
; CURRENT APPLICATION NUMBER: US/10/029,217A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/257,761
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Mus musculus
```

US-10-029-217A-24

Query Match 78.4%; Score 40; DB 9; Length 3907;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
Db 439 CCGAGTCAGC 448

RESULT 10

US-09-804-898-2

; Sequence 2, Application US/09804898
; Patent No. US20020045264A1
; GENERAL INFORMATION:
; APPLICANT: DURING, MATTHEW
; APPLICANT: XIAO, WEIDONG
; TITLE OF INVENTION: PRODUCTION OF CHIMERIC CAPSID VECTORS
; FILE REFERENCE: 102182-14
; CURRENT APPLICATION NUMBER: US/09/804,898
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/189,110
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 4679
; TYPE: PRT
; ORGANISM: adeno-associated virus 2
US-09-804-898-2

Query Match 78.4%; Score 40; DB 10; Length 4679;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
Db 3911 CCGTGGCTAC 3920

RESULT 11

US-08-969-137-1

; Sequence 1, Application US/08969137
; Patent No. US20010018207A1
; GENERAL INFORMATION:
; APPLICANT: KANDEL, ERIC
; APPLICANT: MAYFORD, MARK
; TITLE OF INVENTION: DNA REGULATORY ELEMENT FOR THE
; TITLE OF INVENTION: EXPRESSION OF TRANSGENES IN NEURONS OF THE MOUSE FOREBRAIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER AND DUNHAM
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,137
; FILING DATE: 12-NOV-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)2276821

Query Match 76.5%; Score 39; DB 8; Length 41;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
Db 17 CCGTGCACC 26

RESULT 13

US-09-925-301-950
; Sequence 950, Application US/09925301

; TELEFAX: (212)3910525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-969-137-1

Query Match 76.5%; Score 39; DB 8; Length 41;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
Db 17 CCGTGCACC 26

RESULT 12

US-08-969-137-4
; Sequence 4, Application US/08969137
; Patent No. US20010018207A1
; GENERAL INFORMATION:
; APPLICANT: KANDEL, ERIC
; APPLICANT: MAYFORD, MARK
; TITLE OF INVENTION: DNA REGULATORY ELEMENT FOR THE
; TITLE OF INVENTION: EXPRESSION OF TRANSGENES IN NEURONS OF THE MOUSE FOREBRAIN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER AND DUNHAM
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,137
; FILING DATE: 12-NOV-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)2276821
; TELEFAX: (212)3910525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-969-137-4

Query Match 76.5%; Score 39; DB 8; Length 41;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
Db 17 CCGTGCACC 26

```

; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 950
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-925-301-950

```

```

Query Match          76.5%; Score 39; DB 10; Length 169;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      4 CCGXXCXXC 13
      ||| | |
DB      126 CCGGECADC 135

```

RESULT 14

```

US-09-811-045A-1
; Sequence 1, Application US/09811045A
; Patent No. US20020035080A1
; GENERAL INFORMATION:
; APPLICANT: Scott, Robert E.
; TITLE OF INVENTION: cDNA encoding P2P proteins and use of P2P cDNA-
; TITLE OF INVENTION: derived antibodies and antisense reagents
; TITLE OF INVENTION: In determining the proliferative potential of
; TITLE OF INVENTION: normal, abnormal and cancer cells in animals
; TITLE OF INVENTION: and humans
; FILE REFERENCE: D6386D
; CURRENT APPLICATION NUMBER: US/09/811,045A
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: US 08/801,308
; PRIOR FILING DATE: 1997-02-18
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 1
; LENGTH: 1404
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: PEPTIDE
; OTHER INFORMATION: P2P polypeptide
US-09-811-045A-1

```

```

Query Match          76.5%; Score 39; DB 10; Length 1404;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      4 CCGXXCXXC 13
      ||| | |
DB      76 CCGNSSCDEC 85

```

RESULT 15

```

US-09-975-143-47
; Sequence 47, Application US/09975143
; Patent No. US20020155513A1
; GENERAL INFORMATION:
; APPLICANT: HSU, Daniel, K.

```

```

; APPLICANT: LIU, Fu-Tong
; APPLICANT: DOWLING, Christopher, A.
; TITLE OF INVENTION: GALECTIN EXPRESSION IS INDUCED IN
; TITLE OF INVENTION: CIRRHOTIC LIVER AND HEPATOCELLULAR CARCINOMA
; FILE REFERENCE: DANHSU.001C1
; CURRENT APPLICATION NUMBER: US/09/975,143
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08561
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 914
; TYPE: PRT
; ORGANISM: human
; ORGANISM: human
US-09-975-143-47

```

```

Query Match          74.5%; Score 38; DB 9; Length 914;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      4 CCGXXCXXC 13
      ||| | |
DB      230 CCGGAGCACC 239

```

```

Search completed: January 29, 2003, 09:34:06
Job time : 4.34091 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:57 ; Search time 5.17045 Seconds
(without alignments)
73.978 Million cell updates/sec

Title: US-09-580-201A-2

Perfect score: 73
Sequence: 1 NGVCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	58.9	1917	4 US-09-627-650B-5	Sequence 5, Appli
2	43	58.9	1917	4 US-09-436-063C-5	Sequence 5, Appli
3	41.5	56.8	768	1 US-08-454-455-4	Sequence 4, Appli
4	41	56.2	574	1 US-08-049-473-2	Sequence 2, Appli
5	41	56.2	574	1 US-08-312-648-2	Sequence 2, Appli
6	41	56.2	574	5 PCT-US94-04190-2	Sequence 2, Appli
7	39	53.4	2211	4 US-09-738-884-1	Sequence 1, Appli
8	38.5	52.7	26	4 US-09-073-407-9	Sequence 9, Appli
9	38.5	52.7	26	4 US-09-073-407-12	Sequence 12, Appli
10	38	52.1	485	2 US-08-724-394A-8	Sequence 8, Appli
11	38	52.1	801	1 US-07-906-349A-6	Sequence 6, Appli
12	38	52.1	1345	2 US-08-977-767-3	Sequence 3, Appli
13	38	52.1	1417	4 US-08-900-230-3	Sequence 3, Appli
14	38	52.1	2088	4 US-09-548-372D-13	Sequence 13, Appli
15	38	52.1	2088	4 US-09-548-367D-13	Sequence 13, Appli
16	37.5	51.4	26	4 US-09-073-407-11	Sequence 11, Appli
17	37	50.7	1128	4 US-09-627-650B-11	Sequence 11, Appli
18	37	50.7	1128	4 US-09-436-063C-11	Sequence 11, Appli
19	37	50.7	1404	4 US-08-801-308-1	Sequence 1, Appli
20	37	50.7	1652	4 US-09-627-650B-1	Sequence 1, Appli
21	37	50.7	1652	4 US-09-436-063C-1	Sequence 1, Appli
22	37	50.7	2508	4 US-09-627-650B-7	Sequence 7, Appli
23	37	50.7	2508	4 US-09-436-063C-7	Sequence 7, Appli
24	37	50.7	2544	4 US-09-627-650B-3	Sequence 3, Appli
25	37	50.7	2544	4 US-09-436-063C-3	Sequence 3, Appli
26	37	50.7	2601	4 US-09-627-650B-9	Sequence 9, Appli
27	37	50.7	2601	4 US-09-436-063C-9	Sequence 9, Appli

28 36.5 50.0 25 1 US-08-084-848A-4 Sequence 4, Appli
29 36.5 50.0 25 1 US-08-458-499-4 Sequence 4, Appli
30 36.5 50.0 26 4 US-09-073-407-10 Sequence 10, Appli
31 36.5 50.0 1810 5 PCT-US95-11684-4 Sequence 4, Appli
32 36 49.3 18 1 US-08-084-848A-6 Sequence 6, Appli
33 36 49.3 38 1 US-08-458-499-6 Sequence 6, Appli
34 36 49.3 38 1 US-07-609-716-53 Sequence 53, Appli
35 36 49.3 44 4 US-09-046-894-41 Sequence 41, Appli
36 36 49.3 44 4 US-09-046-894-42 Sequence 42, Appli
37 36 49.3 441 2 US-08-491-835-4 Sequence 4, Appli
38 36 49.3 441 3 US-08-946-092A-4 Sequence 4, Appli
39 36 49.3 441 4 US-09-172-062-4 Sequence 4, Appli
40 36 49.3 441 4 US-09-301-520D-4 Sequence 4, Appli
41 36 49.3 441 5 PCT-US94-00685-4 Sequence 4, Appli
42 35.5 48.6 22 1 US-08-423-399B-27 Sequence 27, Appli
43 35.5 48.6 25 1 US-08-084-848A-1 Sequence 1, Appli
44 35.5 48.6 25 1 US-08-084-848A-5 Sequence 5, Appli
45 35.5 48.6 25 1 US-08-458-499-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-627-650B-5
; Sequence 5, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; TITLE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1917
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-5

Query Match 58.9%; Score 43; DB 4; Length 1917;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GYCCGXXLCHXC 13
Db 901 GTCCGTGTCAC 912

RESULT 2
US-09-436-063C-5
; Sequence 5, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; TITLE OF INVENTION: Nematoe Neuromuscular Junction GABA Receptors and
; TITLE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5

```
; LENGTH: 1917
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-5

Query Match      58.9%; Score 43; DB 4; Length 1917;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY  2 GVCVCGXXLCHXC 13
Db  901 GTCCGTGTCAC 912

RESULT 3
US-08-454-455-4
; Sequence 4, Application US/08454455
; Patent No. 5635601
; GENERAL INFORMATION:
; APPLICANT: Moyle, Matthew
; APPLICANT: McLean, John W.
; TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,455
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193989
; FILING DATE: 09-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/004142
; FILING DATE: 13-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670607
; FILING DATE: 14-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: P0699C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-454-455-4

Query Match      56.8%; Score 41.5; DB 1; Length 768;
Best Local Similarity 72.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY  2 GVC-CGXLCXCH 11
Db  524 GVCVCGKCLCH 534

RESULT 4
US-08-049-473-2
; Sequence 2, Application US/08049473
; Patent No. 5386021
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Mishima, Koichi
; APPLICANT: Nightingale, Maria
; APPLICANT: Tsuchiya, Mikako
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; TITLE OF INVENTION: PROTEIN WITH AN ADP-RYBOSYLATION FACTOR DOMAIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,473
; FILING DATE: 19930419
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH050.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 574 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-049-473-2

Query Match      56.2%; Score 41; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

QY  2 GVC-----CGXXLCHXC 13
Db  32 GVCEDVFSIQGDKVPRLLLCGHTVCHDC 59

RESULT 5
US-08-312-648-2
; Sequence 2, Application US/08312648
; Patent No. 5514600
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Mishima, Koichi
; APPLICANT: Nightingale, Maria
; APPLICANT: Tsuchiya, Mikako
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; TITLE OF INVENTION: PROTEIN WITH AN ADP-RYBOSYLATION FACTOR DOMAIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/312,648
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,473
FILING DATE: 19-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael L.
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH050.001DV1
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-312-648-2

Query Match 56.2%; Score 41; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

QY 2 GVC-----CGXXLCHXC 13
||| ||:| |
Db 32 GVCEDVFSLQGDVKVPRLLLCGHTVCHDC 59

RESULT 6
PCT-US94-04190-2
Sequence 2, Application PC/TUS9404190
GENERAL INFORMATION:
APPLICANT: The Government of the United States of America
APPLICANT: as represented by the Secretary, Department
of Health and Human Services
TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
PROTEIN WITH AN ADP-RYBOSYLATION FACTOR DOMAIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04190
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael L.
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH050.001QPC
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-04190-2

Query Match 56.2%; Score 41; DB 5; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

QY 2 GVC-----CGXXLCHXC 13
||| ||:| |
Db 32 GVCEDVFSLQGDVKVPRLLLCGHTVCHDC 59

RESULT 7
US-09-738-884-1
Sequence 1, Application US/09738884
Patent No. 6391606
GENERAL INFORMATION:
APPLICANT: GUEGLER, Karl et al
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: CL000849
CURRENT APPLICATION NUMBER: US/09/738,884
CURRENT FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2211
TYPE: PRT
ORGANISM: Human
US-09-738-884-1

Query Match 53.4%; Score 39; DB 4; Length 2211;
Best Local Similarity 50.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13
||| | |
Db 1168 GCCCGCCACCTC 1179

RESULT 8
US-09-073-407-9
Sequence 9, Application US/09073407
Patent No. 6232065
GENERAL INFORMATION:
APPLICANT: Robinson, Daniel R.
APPLICANT: Kung, Hsing-Jien
TITLE OF INVENTION: ANALYSIS OF GENE FAMILY EXPRESSION
FILE REFERENCE: CASE-03147
CURRENT APPLICATION NUMBER: US/09/073,407
CURRENT FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-073-407-9

Query Match 52.7%; Score 38.5; DB 4; Length 26;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 NGVCCGXXLCHXC 13
||:| | |
Db 15 NGMCCAAA-CRTC 26

RESULT 9
US-09-073-407-12
Sequence 12, Application US/09073407
Patent No. 6232065

;; GENERAL INFORMATION:
;; APPLICANT: Robinson, Daniel R.
;; APPLICANT: Kung, Hsing-Jien
;; TITLE OF INVENTION: ANALYSIS OF GENE FAMILY EXPRESSION
;; FILE REFERENCE: CASE-03147
;; CURRENT APPLICATION NUMBER: US/09/073,407
;; CURRENT FILING DATE: 1998-05-06
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 12
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-073-407-12

Query Match 52.7%; Score 38.5; DB 4; Length 26;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 NGVCCGXXLCHXC 13
||| | |
Db 15 NGMCCATA-CRTC 26

RESULT 10
US-08-724-394A-8
; Sequence 8, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724.394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region

;; LOCATION: 1..485
;; OTHER INFORMATION: /note= "RoRet"
US-08-724-394A-8
Query Match 52.1%; Score 38; DB 2; Length 485;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 CGXXLCHXC 13
||| | | |
Db 31 CGHSYCHLC 39
RESULT 11
US-07-906-349A-6
; Sequence 6, Application US/07906349A
; Patent No. 5434064
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES
; TITLE OF INVENTION: TARGET PROTEINS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906.349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 801 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-906-349A-6

Query Match 52.1%; Score 38; DB 1; Length 801;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
||| | |
Db 339 GCCCGGTACC 350

RESULT 12
US-08-977-767-3
; Sequence 3, Application US/08977767
; Patent No. 5972684
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Greenwald, Sara

APPLICANT: Corley, Neil C.
TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/977,767
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0423 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1345 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1532042
US-08-977-767-3

Query Match 52.1%; Score 38; DB 2; Length 1345;
Best Local Similarity 50.0%; Pred. No. 6.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXLCHXC 13
||| | |
Db 1089 CCGTACGGC 1098

RESULT 13
US-08-900-230-3
; Sequence 3, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230

; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1417 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-900-230-3

Query Match 52.1%; Score 38; DB 4; Length 1417;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13
| | | | |
Db 958 GGC CGCTTCGCC 969

RESULT 14
US-09-548-372D-13
; Sequence 13, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 2088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-13

Query Match 52.1%; Score 38; DB 4; Length 2088;
Best Local Similarity 55.6%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GVC CGXXLC 10
| | | | |
Db 1566 GACC GAAC 1574

RESULT 15
US-09-548-367D-13
; Sequence 13, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF

```
; FILE REFERENCE: 29915/G280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 2088
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-548-367D-13

Query Match      52.1%; Score 38; DB 4; Length 2088;
Best Local Similarity 55.6%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2  GVCCGXXLC 10
      | | | | |
Db      1566  GACCGAAAC 1574

Search completed: January 29, 2003, 09:33:41
Job time : 6.17045 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:57 ; Search time 24.2614 Seconds
(without alignments)
73.978 Million cell updates/sec

Title: US-09-580-201a-12
Perfect score: 325
Sequence: 1 MRCLPVLIIILLTASAPGV.....ILRGILRNGVCCGYKLCHPC 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	19.1	69	1 US-08-137-800-40	Sequence 40, Appl
2	62	19.1	69	1 US-08-477-383-40	Sequence 40, Appl
3	62	19.1	69	1 US-08-487-174-40	Sequence 40, Appl
4	62	19.1	69	1 US-08-480-750-40	Sequence 40, Appl
5	60.5	18.6	469	1 US-08-313-288B-15	Sequence 15, Appl
6	60	18.5	378	1 US-07-723-002C-8	Sequence 8, Appl
7	59.5	18.3	184	4 US-08-149-101A-23	Sequence 23, Appl
8	59.5	18.3	184	5 PCT-US94-12873-23	Sequence 23, Appl
9	58.5	18.0	787	1 US-08-256-938-4	Sequence 4, Appl
10	58.5	18.0	787	2 US-08-797-689-16	Sequence 16, Appl
11	57	17.5	207	3 US-08-974-022-47	Sequence 47, Appl
12	57	17.5	207	4 US-08-795-445A-47	Sequence 47, Appl
13	57	17.5	207	4 US-08-795-447A-47	Sequence 47, Appl
14	57	17.5	207	4 US-08-974-186-47	Sequence 47, Appl
15	57	17.5	207	4 US-08-795-446B-47	Sequence 47, Appl
16	57	17.5	207	4 US-08-706-945D-133	Sequence 133, App
17	57	17.5	325	1 US-08-292-549-2	Sequence 2, Appl
18	57	17.5	325	4 US-09-042-785A-9	Sequence 9, Appl
19	57	17.5	325	5 PCT-US91-02207-2	Sequence 2, Appl
20	56	17.2	174	1 US-08-434-411-2	Sequence 2, Appl
21	56	17.2	174	1 US-08-434-402-2	Sequence 2, Appl
22	56	17.2	174	1 US-08-783-288-2	Sequence 2, Appl
23	56	17.2	174	2 US-08-890-640-2	Sequence 2, Appl
24	56	17.2	174	3 US-08-833-167-49	Sequence 49, Appl
25	56	17.2	174	3 US-08-833-167-50	Sequence 50, Appl
26	56	17.2	174	3 US-08-833-167-51	Sequence 51, Appl
27	56	17.2	174	3 US-08-833-167-52	Sequence 52, Appl

28	56	17.2	174	3 US-08-833-167-53	Sequence 53, Appl
29	56	17.2	174	3 US-08-833-167-54	Sequence 54, Appl
30	56	17.2	174	3 US-08-833-167-55	Sequence 55, Appl
31	56	17.2	174	3 US-08-833-167-56	Sequence 56, Appl
32	56	17.2	174	3 US-08-833-167-95	Sequence 95, Appl
33	56	17.2	174	3 US-08-833-167-96	Sequence 96, Appl
34	56	17.2	174	3 US-08-833-167-97	Sequence 97, Appl
35	56	17.2	174	3 US-08-833-167-98	Sequence 98, Appl
36	56	17.2	174	3 US-08-833-167-99	Sequence 99, Appl
37	56	17.2	174	3 US-08-833-167-100	Sequence 100, App
38	56	17.2	174	3 US-08-833-167-101	Sequence 101, App
39	56	17.2	174	3 US-08-833-167-102	Sequence 102, App
40	56	17.2	174	3 US-08-833-167-103	Sequence 103, App
41	56	17.2	174	4 US-09-344-837A-49	Sequence 49, Appl
42	56	17.2	174	4 US-09-344-837A-50	Sequence 50, Appl
43	56	17.2	174	4 US-09-344-837A-51	Sequence 51, Appl
44	56	17.2	174	4 US-09-344-837A-52	Sequence 52, Appl
45	56	17.2	174	4 US-09-344-837A-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-08-137-800-40
; Sequence 40, Application US/08137800
; Patent No. 5514774
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hilliard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Santos, Ameurina D.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1000
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,800
; FILING DATE: 19-OCT-1993
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Innen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-104763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus striatus
US-08-137-800-40

Query Match 19.1%; Score 62; DB 1; Length 69;
Best Local Similarity 28.3%; Pred. No. 0.84;
Matches 17; Conservative 11; Mismatches 26; Indels 6; Gaps 3;

QY 6 VLIILLLTASAPGVVLPKTE-DDVPMSSVYNGKSLRGLRNGVCCGY---KLCGPC 61
:: : || | || | : : : ||| : ||| : |||
Db 5 MMFTVLLVLTATNVSTSDRASDGRNAVAHEROKSLVPSVIT--TCCGYDPGTMCPPC 62

RESULT 2

US-08-477-383-40
; Sequence 40, Application US/08477383
; Patent No. 5589340
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,383
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELEPHONE: 202-962-8300
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus striatus

US-08-477-383-40

Query Match 19.1%; Score 62; DB 1; Length 69;
Best Local Similarity 28.3%; Pred. No. 0.84;
Matches 17; Conservative 11; Mismatches 26; Indels 6; Gaps 3;

QY 6 VLIILLLTASAPGVVLPKTE-DDVPMSSVYNGKSLRGLRNGVCCGY---KLCGPC 61
:: : || | || | : : : ||| : ||| : |||
Db 5 MMFTVLLVLTATNVSTSDRASDGRNAVAHEROKSLVPSVIT--TCCGYDPGTMCPPC 62

RESULT 3

US-08-487-174-40
; Sequence 40, Application US/08487174
; Patent No. 5595972
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.

; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,174
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/137,800
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/084,848
; FILING DATE: 29-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24260-107673
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Conus striatus

US-08-487-174-40

Query Match 19.1%; Score 62; DB 1; Length 69;
Best Local Similarity 28.3%; Pred. No. 0.84;
Matches 17; Conservative 11; Mismatches 26; Indels 6; Gaps 3;

QY 6 VLIILLLTASAPGVVLPKTE-DDVPMSSVYNGKSLRGLRNGVCCGY---KLCGPC 61
:: : || | || | : : : ||| : ||| : |||
Db 5 MMFTVLLVLTATNVSTSDRASDGRNAVAHEROKSLVPSVIT--TCCGYDPGTMCPPC 62

RESULT 4

US-08-480-750-40
; Sequence 40, Application US/08480750
; Patent No. 5633347
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.

; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Hillyard, David R.
; APPLICANT: Macintosh, J. Michael
; APPLICANT: Santos, Ameurfino S.
; TITLE OF INVENTION: Conotoxin Peptides
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC

COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480.750
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137.800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084.848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Innen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus striatus
US-08-480-750-40

Query Match 19.1%; Score 62; DB 1; Length 69;
Best Local Similarity 28.3%; Pred. No. 0.84;
Matches 17; Conservative 11; Mismatches 26; Indels 6; Gaps 3;

QY 6 VLIIILLTASAPGVVLPKTE-DQVPMSSVYGNKSKILRGILNGVCGY---KLCGPC 61
DB 5 MMFTVLLVLTANNVSTPSDRASDGRNAVHEROKSLVPSVIT--TCGGYDPGTMCPC 62

RESULT 5
US-08-313-288B-15
; Sequence 15, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and AviHu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313.288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-313-288B-15

Query Match 18.6%; Score 60.5; DB 1; Length 469;
Best Local Similarity 29.2%; Pred. No. 13;
Matches 21; Conservative 7; Mismatches 21; Indels 23; Gaps 4;

QY 4 LPVLIILLTASAPGVVLPKTEDDVPMSVYGNKSKILRGILNGV---CC-----G 54
DB 13 LPPLILLTLTPATG-----SDPVLCTQYESSGCKGLLGGVSVDECCCLNTAFA 63

QY 55 YK-----LCGPC 61
DB 64 YKRGSGGLCQPC 75

RESULT 6
US-07-723-002C-8
; Sequence 8, Application US/07723002C
; Patent No. 5447862
; GENERAL INFORMATION:
; APPLICANT: Heim, Jutta
; APPLICANT: Meyhack, Bernd
; APPLICANT: Gysler, Christof
; APPLICANT: Visser, Jacob
; APPLICANT: Kester, Hermanus Cornelis Maria
; TITLE OF INVENTION: No. 5447862el Expression System
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/723.002C
; FILING DATE: 28-JUN-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8702475
; FILING DATE: 04-FEB-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 150.880
; FILING DATE: 29-JAN-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8818046.8
; FILING DATE: 28-JUL-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8914666.6
; FILING DATE: 26-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 384.898
; FILING DATE: 24-JUL-1989
; ATTORNEY/AGENT INFORMATION:

Query Match 18.3%; Score 59.5; DB 4; Length 184;

```

RESULT 9
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534

```



```
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-938-4

Query Match      18.0%; Score 58.5; DB 1; Length 787;
Best Local Similarity 29.9%; Pred. No. 44;
Matches 20; Conservative 8; Mismatches 20; Indels 19; Gaps 3;

QY 8 IILLLTASAPGVVLPKT-----EDVPMS-----SVYGNKSLIRGLRNGVCC 53
Db 7 ISLFLFSAYSRGVRFTPLGPASSLPQSFLKCLEQVRKIQDGA-----LQEKLC 61

QY 54 GYKLCHP 60
Db 62 TYKLCHP 68

RESULT 10
US-08-797-689-16
; Sequence 16, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 787 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-16

Query Match      18.0%; Score 58.5; DB 2; Length 787;
Best Local Similarity 29.9%; Pred. No. 44;
Matches 20; Conservative 8; Mismatches 20; Indels 19; Gaps 3;

QY 8 IILLLTASAPGVVLPKT-----EDVPMS-----SVYGNKSLIRGLRNGVCC 53
Db 7 ISLFLFSAYSRGVRFTPLGPASSLPQSFLKCLEQVRKIQDGA-----LQEKLC 61

QY 54 GYKLCHP 60
Db 62 TYKLCHP 68

RESULT 11
US-08-974-022-47
; Sequence 47, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-022-47

Query Match      17.5%; Score 57; DB 3; Length 207;
Best Local Similarity 35.2%; Pred. No. 13;
Matches 19; Conservative 4; Mismatches 19; Indels 12; Gaps 3;

QY 7 LIILLLTASAPGVVLPKTEDVPMSVYGNKSLIRGLRNGVCCGYKLCHP 60
Db 4 LIALLVYVYVG-----DDVPYS--NQGKCGHDYKDLCCA--SCHP 45

RESULT 12
US-08-795-445A-47
; Sequence 47, Application US/08795445A
; Patent No. 6284485
```


RESULT 15
US-08-795-446B-47
; Sequence 47, Application US/08795446B
; Patent No. 6288032
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,446B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-795-446B-47

Query Match 17.5%; Score 57; DB 4; Length 207;
Best Local Similarity 35.2%; Pred. No. 13;
Matches 19; Conservative 4; Mismatches 19; Indels 12; Gaps 3;
QY 7 LITLLLTASAPGVVLPKTEDDVPMSVYGNCKSLRGLRNGVCCGYKLCHP 60
|| || | |||| || || :|| || ||
Db 4 LIALVCVVYVG-----DDVPYSS--NCKGCGHDYKDLCCA--SCHP 45

Search completed: January 29, 2003, 09:33:43
Job time : 26.2614 secs

THIS PAGE BLANK (uspro)

Search completed: January 29, 2003, 09:34:08
Job time : 15.5568 secs

Db 22 VAKIQDGAA-----IQEKICATYKICHP 45

RESULT 11

US-09-876-527-16

Sequence 16, Application US/09876527

Patent No. US20020102616a1

GENERAL INFORMATION:

APPLICANT: Kindvogel, Wayne
Jellinek, Laura J.
Sheppard, Paul O.
Hagopian, William A.
Lagasse, James M.

TITLE OF INVENTION: ISLET CELL ANTIGEN 1851

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Zymogenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/876,527

FILING DATE: 07-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/811,481

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Lingenfelter, Susan

REGISTRATION NUMBER: P-41,156

REFERENCE/DOCKET NUMBER: 95-36

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-442-6675

TELEFAX: 206-442-6678

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-09-876-527-16

Query Match 17.1%; Score 55.5; DB 10; Length 1012;

Best local Similarity 31.0%; Pred. No. 1.3e+02;

Matches 18; Conservative 9; Mismatches 18; Indels 13; Gaps 3;

QY 4 LPLVLLILITAPGVVLPKTEDDVPMSSVYGNKSI---LRGILRNVCYCKKIC 58

Db 2 LPLVLLILITAPGVVLPKTEDDVPMSSVYGNKSI---LRGILRNVCYCKKIC 49

RESULT 12

US-09-754-532-94

Sequence 94, Application US/09754532

Patent No. US20010016191a1

GENERAL INFORMATION:

APPLICANT: Osslund, Timothy D.

TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 110

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.
STREET: Amgen Center, 1840 DeHavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: United States of America

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/754,532

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/448,716

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pessin, Karol

REGISTRATION NUMBER: 34,899

TELECOMMUNICATION INFORMATION:

TELEPHONE: 805/499-5725

TELEFAX: 805/499-8011

INFORMATION FOR SEQ ID NO: 94:

SEQUENCE CHARACTERISTICS:

LENGTH: 175 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-754-532-94

Query Match 16.9%; Score 55; DB 10; Length 175;

Best local Similarity 30.4%; Pred. No. 22;

Matches 14; Conservative 8; Mismatches 16; Indels 8; Gaps 2;

QY 15 ASAPGVVLPKTEDDVPMSSVYGNKSIILRGILRNVCYCKKICHP 60

Db 8 SSIPOSFLKALRQ---VRKIQDGAA-----IQEKICATYKICHP 45

RESULT 13

US-09-754-532-95

Sequence 95, Application US/09754532

Patent No. US20010016191a1

GENERAL INFORMATION:

APPLICANT: Osslund, Timothy D.

TITLE OF INVENTION: G-CSF ANALOG COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 110

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Inc.
STREET: Amgen Center, 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California

COUNTRY: United States of America

ZIP: 91320-1789

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/754,532

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/448,716

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Pessin, Karol

REGISTRATION NUMBER: 34,899

TELECOMMUNICATION INFORMATION:

TELEPHONE: 805/499-5725

TELEFAX: 805/499-8011

INFORMATION FOR SEQ ID NO: 95:

Db 61 CG--KCQC 67

RESULT 5

US-09-811-672-10

; Sequence 10, Application US/09811672

; Patent No. US20020052490A1

; GENERAL INFORMATION:

; APPLICANT: BALL, Tanja

; APPLICANT: VRTALA, Susanne

; APPLICANT: SPERR, Wolfgang

; APPLICANT: VALENT, Peter

; APPLICANT: SUSANI, Markus

; APPLICANT: KRAFT, Dietrich

; APPLICANT: VALENTA, Rudolf

; APPLICANT: LAFFER, Sylvia

; TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF, CORRESPONDING RECOMBINANT MOLECULES, VECTORS AND HOSTS CONTAINING THE DNA MOLECULES, DIAGNOSTIC USES OF SAID ALLERGENS AND FRAGMENTS

; FILE REFERENCE: 1614-0247P

; CURRENT APPLICATION NUMBER: US/09/811,672

; CURRENT FILING DATE: 2001-03-20

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 10

; LENGTH: 263

; TYPE: PRT

; ORGANISM: Timothy Grass

US-09-811-672-10

Query Match 18.0%; Score 58.5; DB 10; Length 263;

Best Local Similarity 28.6%; Pred. No. 14;

Matches 18; Conservative 9; Mismatches 23; Indels 13; Gaps 2;

Qy 4 LPVLILLLLLTASAPGVVLPKTEDDVPMSVYGN-----GKSILRGILRNGVCC 53

Db 8 LLVVVLFVFLGSAYGI---PKVPPGNITATYGDKWLDAKSTWYGKPTGAGPKDNGGAC 64

Qy 54 GYK 56

Db 65 GYK 67

RESULT 6

US-09-984-186-16

; Sequence 16, Application US/09984186

; Patent No. US2002015101A1

; GENERAL INFORMATION:

; APPLICANT: Fleer, Reinhard

; Fournier, Alain

; Guitton, Jean-Dominique

; Jung, Gerard

; Yeh, Patrice

; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION CONTAINING SAID POLYPEPTIDES

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Rhone-Poulenc Rorer Inc.

; STREET: 500 Arcola Road, 3C43

; CITY: Collegeville

; STATE: PA

; COUNTRY: USA

; ZIP: 19426

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Macintosh

; OPERATING SYSTEM: System 7.1

; SOFTWARE: Word 5.1 (PatentIn)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/984,186

; FILING DATE: 29-Oct-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/797,689

; FILING DATE: 31-JAN-1997

; APPLICATION NUMBER: US 08/256,927

; FILING DATE: 28-JUL-1994

; APPLICATION NUMBER: FR 92/01064

; FILING DATE: 31-JAN-1992

; APPLICATION NUMBER: PCT/FR93/00085

; FILING DATE: 28-JAN-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith Ph.D., Julie K.

; REGISTRATION NUMBER: P-38,619

; REFERENCE/DOCKET NUMBER: ST92006-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (610) 454-3839

; TELEFAX: (610) 454-3808

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 787 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-09-984-186-16

Query Match 18.0%; Score 58.5; DB 10; Length 787;

Best Local Similarity 29.9%; Pred. No. 46;

Matches 20; Conservative 8; Mismatches 20; Indels 19; Gaps 3;

Qy 8 IILLLLLTASAPGVVLPKTEDDVPMS-----SVYNGKSIILRGILRNGVCC 53

Db 7 ISLLFLFSAYSRGVVRTPLGPASSLPQSFLKLCLEQVRKIQDGAA-----LQEKLC 61

Qy 54 GYKICHP 60

Db 62 TYKLCHP 68

RESULT 7

US-09-894-882-329

; Sequence 329, Application US/09894882

; Patent No. US20020102607A1

; GENERAL INFORMATION:

; APPLICANT: University of Utah Research Foundation

; APPLICANT: Cognetix Inc.

; APPLICANT: Walker, Craig S.

; APPLICANT: Shetty, Reshma

; APPLICANT: Jimenez, Elsie C.

; APPLICANT: McIntosh, J. Michael

; APPLICANT: Olivera, Baldomero M.

; APPLICANT: Watkins, Maren

; APPLICANT: Jones, Robert M.

; APPLICANT: Shen, Greg S.

; TITLE OF INVENTION: I-Superfamily Conotoxins

; FILE REFERENCE: 2314-238

; CURRENT APPLICATION NUMBER: US/09/894,882

; CURRENT FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 60/

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 60/243,410

; PRIOR FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: US 60/246,581

; PRIOR FILING DATE: 2000-11-08

; PRIOR APPLICATION NUMBER: US 60/247,714

; PRIOR FILING DATE: 2000-11-14

; PRIOR APPLICATION NUMBER: US 60/264,256

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 506

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 329

; LENGTH: 72

; TYPE: PRT

; ORGANISM: Conus episcopus


```
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 6787
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6787

Query Match          19.5%; Score 63.5; DB 9; Length 625;
Best Local Similarity 36.2%; Pred. No. 9,7;
Matches 17; Conservative 8; Mismatches 17; Indels 5; Gaps 1;

QY 4 LPLVILLLLTASAPGVVVLPTKEDDVPMSVYGNKSGILRNG 50
      | | | | | | | | | | | | | | | | | | | | |
DB 246 LNVELALTLLVIACPGALVI-----SIPSVIVAGISAKDGLIKG 287

RESULT 3
US-09-864-761-38941
; Sequence 38941, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38941
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005661.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.7
; OTHER INFORMATION: SWISSPROT HIT: P39584, EVALUOE 1.70e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF312621.1, EVALUOE 1.40e+00
US-09-864-761-38941

Query Match          18.9%; Score 61.5; DB 10; Length 38;
Best Local Similarity 36.7%; Pred. No. 0,75;
Matches 18; Conservative 4; Mismatches 8; Indels 19; Gaps 3;

QY 18 PGVVVLPTKEDDVPMSVYGNKSGILRNGVCCGYKLCH 59
      | | | | | | | | | | | | | | | | | | | | |
DB 2 PQEMLPR-----GPGKALPWLPLLLSLDGLRPGCCGL-LCH 38

RESULT 4
US-09-950-933A-42
; Sequence 42, Application US/09950933A
; Patent No. US20020166141A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro, Pedro
; TITLE OF INVENTION: Antimicrobial Peptides and Methods of
; FILE REFERENCE: 35718/238472
; CURRENT APPLICATION NUMBER: US/09/950,933A
; CURRENT FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/232,569
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Zea mays
US-09-950-933A-42

Query Match          18.8%; Score 61; DB 9; Length 96;
Best Local Similarity 27.5%; Pred. No. 2,4;
Matches 19; Conservative 11; Mismatches 29; Indels 10; Gaps 2;

QY 1 MRCLEVLIIILLTASAPGVVVLPTKEDDVPMSVYGNKSGILRNG-----VC 52
      | | | | | | | | | | | | | | | | | | | | |
DB 1 MKAIPVALLLLVLVAASFFKHLAEADGAVPDGVCDCGKCRSRCSLKKAGRCMGLCMWC 60

QY 53 CGYKLCCHPC 61
      | | | | |
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:29:32 ; Search time 14.5568 Seconds
(without alignments)
84.558 Million cell updates/sec

Title: US-09-580-201A-12
Perfect score: 325
Sequence: 1 MRCPLVLIILLTASAPGV.....ILRGLRNGVCCGKYLCHPC 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	64	19.5	96	12	US-10-016-481-11
2	63.5	19.5	625	9	US-09-738-626-6787
3	61.5	18.9	38	10	US-09-864-761-38941
4	61	18.8	96	9	US-09-950-933A-42
5	58.5	18.0	263	10	US-09-811-672-10
6	58.5	18.0	787	10	US-09-884-186-16
7	57.5	17.7	72	10	US-09-894-882-329
8	57.5	17.7	72	10	US-09-894-882-341
9	57.5	17.7	625	9	US-09-738-626-4801
10	55.5	17.1	175	10	US-09-754-532-86
11	55.5	17.1	1012	10	US-09-876-527-16
12	55	16.9	175	10	US-09-754-532-94
13	55	16.9	175	10	US-09-754-532-95
14	55	16.9	175	10	US-09-754-532-96
15	55	16.9	175	10	US-09-754-532-99
16	55	16.9	495	10	US-09-925-301-1312
17	54.5	16.8	175	10	US-09-754-532-87
18	54.5	16.8	436	10	US-09-815-242-10988
19	54.5	16.8	476	10	US-09-815-242-11910

20	54	16.6	429	9	US-10-038-072-94
21	54	16.6	440	10	US-09-815-242-13429
22	53.5	16.5	108	12	US-10-016-481-5
23	53.5	16.5	174	9	US-09-921-114-1
24	53.5	16.5	174	10	US-09-754-532-82
25	53.5	16.5	174	10	US-09-760-008A-1
26	53.5	16.5	174	10	US-09-950-473-2
27	53.5	16.5	174	12	US-10-016-403-1
28	53.5	16.5	174	12	US-10-016-403-2
29	53.5	16.5	174	12	US-10-016-403-3
30	53.5	16.5	174	12	US-10-016-403-4
31	53.5	16.5	174	12	US-10-003-496-1
32	53.5	16.5	175	9	US-09-921-114-2
33	53.5	16.5	175	10	US-09-754-532-2
34	53.5	16.5	175	10	US-09-754-532-67
35	53.5	16.5	175	10	US-09-754-532-68
36	53.5	16.5	175	10	US-09-754-532-69
37	53.5	16.5	175	10	US-09-754-532-71
38	53.5	16.5	175	10	US-09-754-532-76
39	53.5	16.5	175	10	US-09-754-532-79
40	53.5	16.5	175	10	US-09-754-532-80
41	53.5	16.5	175	10	US-09-754-532-81
42	53.5	16.5	175	10	US-09-754-532-85
43	53.5	16.5	175	10	US-09-754-532-88
44	53.5	16.5	175	10	US-09-754-532-90
45	53.5	16.5	175	10	US-09-754-532-91

ALIGNMENTS

RESULT 1
US-10-016-481-11
; Sequence 11, Application US/10016481
; Patent No. US20020115610A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Qun-Yong
; TITLE OF INVENTION: Prokineticin Polypeptides, Related
; TITLE OF INVENTION: Compositions and Methods.
; FILE REFERENCE: P-UC 5016
; CURRENT APPLICATION NUMBER: US/10/016.481
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/245,882
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Bombina variegata
US-10-016-481-11

Query Match 19.7%; Score 64; DB 12; Length 96;
Best Local Similarity 22.1%; Pred. No. 1.1;
Matches 17; Conservative 8; Mismatches 20; Indels 32; Gaps 2;
Qy 1 MRCPLVLIILLTASAPGVVLPKTEDDVPMSVYNGKSLRGLRNGVCCGKYL---- 57
| | :
Db 1 MKCFAQIVLLLVIAFSGAVITGACDKDVQCGS-----GTCCAASAWSR 45

Qy 58 -----CHP 60
| | |
Db 46 NIRCPIPLNGSGEDCHP 62

RESULT 2
US-09-738-626-6787
; Sequence 6787, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	164	50.5	64	5	Q9BP69	Q9bp69	conus penna
2	160.5	49.4	64	5	Q9BP67	Q9bp67	conus ventr
3	149	45.8	63	5	Q9BP69	Q9bp69	conus texti
4	148	45.5	62	5	Q9BP68	Q9bp68	conus penna
5	147	45.2	64	5	Q9BH75	Q9bh75	conus tessu
6	146.5	45.1	63	5	Q9BP64	Q9bp64	conus penna
7	144	44.3	63	5	Q9BP67	Q9bp67	conus penna
8	143	44.0	64	5	Q9BP61	Q9bp61	conus penna
9	142.5	43.8	62	5	Q9BH86	Q9bh86	conus penna
10	142.5	43.8	63	5	Q9BP68	Q9bp68	conus penna
11	142	43.7	61	5	Q9BP62	Q9bp62	conus ventr
12	141	43.4	62	5	Q9BP60	Q9bp60	conus penna
13	138	42.5	64	5	Q9BP63	Q9bp63	conus penna
14	138	42.5	65	5	Q9BP65	Q9bp65	conus tessu
15	136	41.8	64	5	Q9BP66	Q9bp66	conus tessu
16	134	41.2	64	5	Q9BP64	Q9bp64	conus tessu

Qy	1	MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSKILRGILNGVCC--GYKLCH	58
Dd	1	MRCPLVFVILLITLASAPSDARPDKDIPLVSFQDNAKRALQILSNKRYYCCYFDSCC	60
 RESULT 9			
Q9BH86		PRELIMINARY; PRT; 62 AA.	
ID	Q9BH86		
AC	Q9BH86;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	Conotoxin scaffold IX precursor.		
OS	Conus pennaceus (Feathered cone).		
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;		
OX	Neogastropoda; Conoidea; Conidae; Conus.		
NCBI_TaxID=37335;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21105969; PubMed=11158371;		
RA	Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,		
RA	Fainzilber M.;		
RT	"Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";		
RL	Mol. Biol. Evol. 18:120-131(2001).		
DR	EMBL; AF215098; AAG60519.1; -		
DR	EMBL; AF214954; AAG60382.1; -		
SQ	SEQUENCE 62 AA; 7127 MW; 45D7118669BD051C CRC64;		
 Query Match 43.8%; Score 142.5; DB 5; Length 62;			
Best Local Similarity 50.0%; Pred. No. 2.2e-10;			
Matches 30; Conservative 8; Mismatches 21; Indels 1; Gaps 1;			
Qy	1	MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSKILRGILNGVCC-GYKLCH	59
Dd	1	MRCPLVFVILLITLASAPSDARPDKDIPLVSFQDHAKRILOTFESRYDCCKTFCCH	60
 RESULT 10			
Q9BPF8		PRELIMINARY; PRT; 63 AA.	
ID	Q9BPF8		
AC	Q9BPF8;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	Conotoxin scaffold IX.		
OS	Conus pennaceus (Feathered cone).		
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;		
OX	Neogastropoda; Conoidea; Conidae; Conus.		
NCBI_TaxID=37335;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21105969; PubMed=11158371;		
RA	Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,		
RA	Fainzilber M.;		
RT	"Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";		
RL	Mol. Biol. Evol. 18:120-131(2001).		
DR	EMBL; AF214971; AAG60399.1; -		
SQ	SEQUENCE 63 AA; 6878 MW; 27DD9A6AD056B548 CRC64;		
 Query Match 43.8%; Score 142.5; DB 5; Length 63;			
Best Local Similarity 52.5%; Pred. No. 2.2e-10;			
Matches 32; Conservative 8; Mismatches 20; Indels 1; Gaps 1;			
Qy	1	MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSKILRGILNGVCC-GYKLCH	59
Dd	1	MRCPLVFVILLITLASAPSDARPDKDIPLVSFQDHLQRTVTLTDIRMCLGTSGCC	60
Qy	60 P 60		
Dd	61 P 61		
 RESULT 11			


```
RP SEQUENCE FROM N.A.
RX MEDLINE-21105969; PubMed-11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RT Fainzilber M.;
  "Mechanisms for Evolving Hypervariability: The Case of Conopeptides.";
RL Mol. Biol. Evol. 18:120-131(2001).
DR EMBL: AF214984; AAG60412.1; -.
DR InterPro; IPR001230; Prenyl_site.
DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
SQ SEQUENCE 64 AA; 7005 MW; 7BD3F180FC09EF11 CRC64;

  Query Match          41.8%; Score 136; DB 5; Length 64;
  Best Local Similarity 51.7%; Pred.No. 1.4e-09;
  Matches 31; Conservative 5; Mismatches 22; Indels 2; Gaps 1;

Qy 1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSVYG--NGKSILRGILRNGVCCGYKLC 58
   | | | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MHCLPVPVILLIIASTPSVDARPKTKDDVPPASPHGADNANRIILTLWNLRGCCEDKTC 60

Search completed: January 29, 2003, 09:32:17
Job time : 48.8295 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 13.8636 Seconds
(without alignments)
182.496 Million cell updates/sec

Title: US-09-580-201A-12
Perfect score: 325
Sequence: 1 MRCLPVLIIILLTASAPGV.....ILRGLRNGVCCGYKLCHPC 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	325	100.0	61	1 CXL2_CONMR	P58808 conus marmo
2	148.5	45.7	62	1 CX52_CONGL	Q9u628 conus glori
3	140	43.1	63	1 CX51_CONGL	Q9u629 conus glori
4	138	42.5	62	1 CX51_CONTE	Q9u700 conus texti
5	136	41.8	63	1 CX5A_CONPU	Q9u626 conus purpu
6	127	39.1	67	1 CXET_CONTE	R81755 conus texti
7	104	32.0	62	1 CX51_CONIM	Q9u625 conus imper
8	83	25.5	13	1 CXL4_CONMR	P58810 conus marmo
9	77	23.7	11	1 CXL1_CONMR	P58807 conus marmo
10	64	19.7	96	1 BV8_BOWVA	Q9pw66 bombina var
11	61	18.8	249	1 MCT1_CANFA	P21842 canis famil
12	60.5	18.6	112	1 FER_PSEAE	Q51383 pseudomonas
13	60.5	18.6	469	1 PROP_HUMAN	P27918 homo sapien
14	60	18.5	1324	1 MSB6_ARATH	O04716 arabidopsis
15	59.5	18.3	110	1 FER_ECOLI	P25528 escherichia
16	59	18.2	12	1 CXL3_CONMR	P58809 conus marmo
17	59	18.2	381	1 CYCH_RHEIM	P45400 rhizobium m
18	59	18.2	410	1 SAHH_THEVO	Q97924 thermoplas
19	58.5	18.0	221	1 CYNH_HELPJ	Q9zn54 helicobacte
20	58.5	18.0	263	1 MP11_PHLPR	P43213 phleum prat
21	57.5	17.7	106	1 INS2_XENLA	P12707 xenopus lae
22	57	17.5	325	1 VT2_SERVA	P25943 shope fibro
23	57	17.5	466	1 IMP2_DROME	P29681 drosophila
24	57	17.5	1353	1 VGL2_CVHOC	P36334 human coron
25	56.5	17.4	269	1 MP11_PRAAQ	Q41260 phalaris aq
26	56	17.2	243	1 NGF_BUNMU	P34128 bungarus mu
27	56	17.2	397	1 YH28_HAEIN	O05087 haemophilus
28	56	17.2	407	1 CARA_BRUME	Q8yib8 bruceella me
29	56	17.2	475	1 PPB_SERMA	P19147 serratia ma
30	56	17.2	969	1 PAC4_HUMAN	P29122 homo sapien
31	55.5	17.1	221	1 CYNH_HELPY	O24855 helicobacte
32	55.5	17.1	359	1 RRP4_YEAST	P38792 saccharomyc
33	55.5	17.1	461	1 PPB4_BACSU	P19406 bacillus su

RESULT 1				
CXL2_CONMR				
ID	CXL2_CONMR	STANDARD;	PRT;	61 AA.
AC	P58808;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lambda-conotoxin CMrVIB precursor (Chi-conotoxin Mria) (Chi-Mria)			
DE	(mr10a).			
OS	Conus marmoreus (Marble cone).			
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;			
OC	Neogastropoda; Conoidea; Conidae; Conus.			
OX	NCBI_TaxID=42752;			
RN	[1]			
RP	SEQUENCE FROM N.A., SEQUENCE OF 49-61, SYNTHESIS, AND MASS			
RP	SPECTROMETRY.			
RC	TISSUE=Venom duct, and Venom;			
RX	MEDLINE=20490660; PubMed=10900201;			
RA	McIntosh J.M., Corpus G.O., Layer R.T., Garrett J.E., Wagstaff J.D.,			
RA	Bulaj G., Vyazovkina A., Yoshikami D., Cruz L.J., Olivera B.M.;			
RT	"Isolation and Characterization of a novel conus peptide with apparent			
RT	antinoceptive activity".			
RL	J. Biol. Chem. 275:32391-32397(2000).			
RN	[2]			
RP	SEQUENCE OF 49-61, AND MASS SPECTROMETRY.			
RC	TISSUE=Venom;			
RX	MEDLINE=20564325; PubMed=10988292;			
RA	Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,			
RA	Seow K.T., Bay B.-H.;			
RT	"Lambda-conotoxins, a new family of conotoxins with unique disulfide			
RT	pattern and protein folding. Isolation and characterization from the			
RT	venom of Conus marmoreus.;"			
RL	J. Biol. Chem. 275:39516-39522(2000).			
RN	[3]			
RP	SEQUENCE OF 49-61, SYNTHESIS, MASS SPECTROMETRY, AND STRUCTURE BY NMR.			
RC	TISSUE=Venom;			
RX	MEDLINE=21419681; PubMed=11528421;			
RA	Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,			
RA	Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.;			
RA	Lewis R.J.;			
RT	"Two new classes of conopeptides inhibit the alpha1-adrenoceptor and			
RT	noradrenaline transporter.;"			
RL	Nat. Neurosci. 4:902-907(2001).			
CC	-!- FUNCTION: Inhibits the neuronal noradrenaline transporter.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- TISSUE SPECIFICITY: Expressed by the venom duct.			
CC	-!- PTM: Exists in two forms, due to cis-trans isomerization at His-			
CC	59-Hyp-60.			
CC	-!- MASS SPECTROMETRY: MW=1408.5; METHOD=Electrospray.			
CC	-!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.			
KW	Neurotoxin; Toxin; Hydroxylation; Signal.			
FT	SIGNAL	1	19	POTENTIAL.
FT	PROPEP	20	48	
FT	PEPTIDE	49	61	LAMBDA-CONOTOXIN CMRVIB.
FT	DISULFID	52	61	
FT	DISULFID	53	58	

34	55.5	17.1	1013	1	PTPX_MACNE	O02695 macaca neme
35	55	16.9	55	1	M84C_DROME	Q01644 drosophila
36	55	16.9	220	1	IF6_PYRAE	Q82t1l pyrobaculum
37	55	16.9	278	1	P29K_STRGC	P42361 streptococc
38	55	16.9	312	1	OXYR_STRVD	Q9x5p2 streptomyc
39	55	16.9	431	1	NORF_CHLMU	Q9p1i3 chlamydia m
40	55	16.9	431	1	NORF_CHLTR	O84745 chlamydia t
41	55	16.9	792	1	SYFB_CHLPN	Q947w0 chlamydia p
42	55	16.9	2245	1	MYSJ_DICDI	P54697 dictyosteli
43	54.5	16.8	208	1	CSF3_MOUSE	P09920 mus musculu
44	54.5	16.8	273	1	HIS6_METAC	Q8t96 methanosarc
45	54.5	16.8	354	1	RFC5_YEAST	P38251 saccharomyc

ALIGNMENTS

```

FT  MOD_RES      60      60      HYDROXYLATION.
SQ  SEQUENCE     61 AA;  6499 MW;  F4DE5B5A97EB8DBA CRC64;

Query Match      100.0%; Score 325; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 5.7e-31;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSLIRLNGVCCGYKLCHP 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  1 MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSLIRLNGVCCGYKLCHP 60

QY  61 C 61
    |
Db  61 C 61

RESULT 2
CX52_CONGL      STANDARD;          PRT;      62 AA.
ID  CX52_CONGL
AC  Q9U6Z8;
DT  15-JUN-2002 (Rel. 41, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Putative conotoxin Gm5.2 precursor.
OS  Conus gloriamaris (Glory of the sea).
OC  Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OX  NCBI_TaxID=37336;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Venom duct;
RX  MEDLINE=99452958; PubMed=10521453;
RA  Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,
RA  Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA  Bandyopadhyay P., Craig A.G., Olivera B.M.;
RT  "The T-superfamily of conotoxins.";
RL  J. Biol. Chem. 274:30664-30671(1999).
RN  [2]
RP  ERRATUM.
RA  Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,
RA  Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA  Bandyopadhyay P., Craig A.G., Olivera B.M.;
RL  J. Biol. Chem. 274:30630-36030(1999).
CC  -!- FUNCTION: Not known. Has only been studied at nucleotidic level.
CC  -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC  -!- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF167166; AAF03686.1; -.
KW  Toxin; Signal.
FT  SIGNAL          1      22      POTENTIAL.
FT  PROPEP         23      49      POTENTIAL.
FT  PEPTIDE        50      60      PUTATIVE CONOTOXIN GM5.2.
FT  DISULFID       51      58      POTENTIAL.
FT  DISULFID       52      59      POTENTIAL.
FT  MOD_RES        60      60      AMIDATION (G-61 PROVIDE AMIDE GROUP)
SQ  SEQUENCE     62 AA;  6817 MW;  6C8C0ADB810732C9 CRC64;

Query Match      45.7%; Score 148.5; DB 1; Length 62;
Best Local Similarity 58.5%; Pred. No. 1.3e-10;
Matches 31; Conservative 8; Mismatches 13; Indels 1; Gaps 1;

QY  1 MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSLIRLNGVCC 53
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

Db  1 MRCPLVFIILLIIASAPSDAQPKTKDDVPLAPLHDNIRSTLQ-TLRKKVCC 52

RESULT 3
CX51_CONGL      STANDARD;          PRT;      63 AA.
ID  CX51_CONGL
AC  Q9U6Z9;
DT  15-JUN-2002 (Rel. 41, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Putative conotoxin Gm5.1 precursor.
OS  Conus gloriamaris (Glory of the sea).
OC  Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OX  NCBI_TaxID=37336;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Venom duct;
RX  MEDLINE=99452958; PubMed=10521453;
RA  Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,
RA  Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA  Bandyopadhyay P., Craig A.G., Olivera B.M.;
RT  "The T-superfamily of conotoxins.";
RL  J. Biol. Chem. 274:30664-30671(1999).
RN  [2]
RP  ERRATUM.
RA  Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,
RA  Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA  Bandyopadhyay P., Craig A.G., Olivera B.M.;
RL  J. Biol. Chem. 274:36030-36030(1999).
CC  -!- FUNCTION: Not known. Has only been studied at nucleotidic level.
CC  -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC  -!- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF167165; AAF03685.1; -.
KW  Toxin; Signal.
FT  SIGNAL          1      21      POTENTIAL.
FT  PROPEP         22      50      POTENTIAL.
FT  PEPTIDE        51      63      PUTATIVE CONOTOXIN GM5.1.
FT  DISULFID       52      59      POTENTIAL.
FT  DISULFID       53      60      POTENTIAL.
SQ  SEQUENCE     63 AA;  7398 MW;  00B24439E5DDE272 CRC64;

Query Match      43.1%; Score 140; DB 1; Length 63;
Best Local Similarity 52.8%; Pred. No. 1.3e-09;
Matches 28; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY  1 MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGNKSLIRLNGVCC 53
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  1 MRCPLVFIILLIIASIFSDTVOLTKDDMPASPHGNGRILRLMLSNKRICC 53

RESULT 4
CX51_CONTE      STANDARD;          PRT;      62 AA.
ID  CX51_CONTE
AC  Q9U700;
DT  15-JUN-2002 (Rel. 41, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Putative conotoxin Tx5.1 precursor.
OS  Conus textile (Cloth-of-gold cone).
OC  Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OX  NCBI_TaxID=6494;

```

```

RA SEQUENCE FROM N.A.
RP TISSUE=Venom duct;
RX MEDLINE=99452958; PubMed=10521453;
RA Walker C.S., Steel D., Jacobsen R.B., Lirazzen M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RT "The T-superfamily of conotoxins.";
RL J. Biol. Chem. 274:30664-30671(1999).
RN [2]
RP ERRATUM.
RA Walker C.S., Steel D., Jacobsen R.B., Lirazzen M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RL J. Biol. Chem. 274:30630-30630(1999).
CC -!- FUNCTION: Not known. Has only been studied at nucleotidic level.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF167164; AAF03684.1; -
KW Toxin; Signal; Amidation.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 49 POTENTIAL.
FT PEPTIDE 50 60 PUTATIVE CONOTOXIN TX5.1.
FT DISULFID 50 57 POTENTIAL.
FT DISULFID 51 58 POTENTIAL.
FT MOD_RES 60 60 AMIDATION (G-61 PROVIDE AMIDE GROUP)
FT (POTENTIAL).
FT SEQUENCE 62 AA: 6857 MW; CE29803DEB0DA21 CRC64;

Query Match 42.5%; Score 138; DB 1; Length 62;
Best Local Similarity 54.7%; Pred. No. 2.2e-09;
Matches 29; Conservative 8; Mismatches 14; Indels 2; Gaps 1;

QY 1 MRCPLPVLIIILLTASAPGVVLPKTEDVPMSSVYNGKSLIRLNGVCC 53
D 1 MRCPLPVLIIILLTASAPGVVLPKTEDVPMSSVYNGKSLIRLNGVCC 51
D 1 MRCPLPVLIIILLTASAPGVVLPKTEDVPMSSVYNGKSLIRLNGVCC 51

RESULT 5
CX5A_CONPU
ID CX5A_CONPU STANDARD; PRT; 63 AA.
AC Q9U6Z6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Conotoxin p5a precursor.
OS Conus purpurascens (Purple cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=41690;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 51-62, MASS SPECTROMETRY, AND
RP SYNTHESIS.
RX TISSUE=Venom duct, and Venom;
RX MEDLINE=99452958; PubMed=10521453;
RA Walker C.S., Steel D., Jacobsen R.B., Lirazzen M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RT "The T-superfamily of conotoxins.";
RL J. Biol. Chem. 274:30664-30671(1999).
RN [2]
RP ERRATUM.
RA Walker C.S., Steel D., Jacobsen R.B., Lirazzen M.B., Cruz L.J.,
```

```

RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RL J. Biol. Chem. 274:30630-30630(1999).
CC -!- FUNCTION: Low levels of the peptide injected into male specimens
CC of the Siamese fighting fish causes an immediate aggressive
CC display in this fish in response to their reflection when placed
CC in a mirrored aquarium; High levels of the peptide suppressed this
CC behavior. No effect is observed when injected into mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1337.5; METHOD=LSIMS.
CC -!- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF167168; AAF03688.1; -
KW Toxin; Amidation; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 47 CONOTOXIN P5A.
FT PEPTIDE 51 62
FT DISULFID 52 59
FT DISULFID 53 60
FT MOD_RES 62 62
FT SEQUENCE 63 AA: 7102 MW; 82A28478C13D7EA9 CRC64;

Query Match 41.8%; Score 136; DB 1; Length 63;
Best Local Similarity 50.9%; Pred. No. 3.8e-09;
Matches 27; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 1 MRCPLPVLIIILLTASAPGVVLPKTEDVPMSSVYNGKSLIRLNGVCC 53
D 1 MRCPLPVLIIILLTASAPGVVLPKTEDVPMSSVYNGKSLIRLNGVCC 53
D 1 MRCPLPVLIIILLTASAPGVVLPKTEDVPMSSVYNGKSLIRLNGVCC 53

RESULT 6
CXET_CONFE
ID CXET_CONFE STANDARD; PRT; 67 AA.
AC P81755; Q9U6Z7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epsilon-conotoxin TXIX precursor (Conotoxin tx5a).
OS Conus textile (Cloth-of-gold cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6494;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 51-63, AND MASS SPECTROMETRY.
RX TISSUE=Venom duct, and Venom;
RX MEDLINE=99452958; PubMed=10521453;
RA Walker C.S., Steel D., Jacobsen R.B., Lirazzen M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RT "The T-superfamily of conotoxins.";
RL J. Biol. Chem. 274:30664-30671(1999).
RN [2]
RP ERRATUM.
RA Walker C.S., Steel D., Jacobsen R.B., Lirazzen M.B., Cruz L.J.,
RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,
RA Bandyopadhyay P., Craig A.G., Olivera B.M.;
RL J. Biol. Chem. 274:30630-30630(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Venom duct;
RX MEDLINE=21105969; PubMed=11158371;
RA Conticello S.G., Gilad Y., Avidan N., Ben-Asher E., Levy Z.,
RA Fainzilber M.;
```

RT "Mechanisms for evolving hypervariability: the case of conopeptides.";

RL Mol. Biol. Evol. 18:120-131(2001).

RN [4]

RP SEQUENCE OF 51-63, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY

RQ NMR.

RC TISSUE=Venom;

RX MEDLINE=99254114; PubMed=10318957;

RA Rigby A.C., Lucas-Munier E., Kalume D.E., Czerwicz E., Hambe B.,

RA Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,

RA Furie B.C., Furie B., Stenflo J.P.;

RT "A conotoxin from *Conus textile* with unusual posttranslational

RT modifications reduces presynaptic Ca²⁺ influx.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).

RN [5]

RP SEQUENCE OF 51-63, AND MASS SPECTROMETRY.

RC TISSUE=Venom;

RX PubMed=10679974;

RA Kalume D.E., Stenflo J.P., Czerwicz E., Hambe B., Furie B.C.,

RA Furie B., Roepstorff P.;

RT "Structure determination of two conotoxins from *Conus textile* by a

RT combination of matrix-assisted laser desorption/ionization

RT time-of-flight and electrospray ionization mass spectrometry and

RT biochemical methods.";

RL J. Mass Spectrom. 35:145-156(2000).

CC -!- FUNCTION: Acts at presynaptic membranes, blocking the calcium

CC channels. Causes hyperactivity upon intracranially injection into

CC mice. Causes dorsal fins drooping in fish.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.

CC -!- PTM: O-glycan consists of the disaccharide Gal-GalNac.

CC -!- MASS SPECTROMETRY: MW=1929.4; METHOD=Electrospray.

CC -!- MASS SPECTROMETRY: MW=1929.4; METHOD=MALDI.

CC -!- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS.

CC TYPE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF167167; AAF03687.1; -.

DR EMBL; AF214958; AAG60386.1; -.

DR PDB; 1WCT; 08-JUN-99.

KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;

KW Vitamin K; Signal; Gamma-carboxyglutamic acid; Glycoprotein;

KW Hydroxylation; Bromination; 3D-structure.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 50 EPSILON-CONOTOXIN TXIX.

FT PEPTIDE 51 63

FT PROPEP 64 67

FT DISULFID 52 58

FT DISULFID 53 59

FT MOD_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.

FT MOD_RES 54 54 GAMMA-CARBOXYGLUTAMIC ACID.

FT MOD_RES 57 57 BROMINATION.

FT MOD_RES 63 63 HYDROXYLATION.

FT CARBOHYD 60 60 O-LINKED (GALNAC...).

SQ SEQUENCE 67 AA; 7587 MW; 7270505504D6BB3D CRC64;

Query Match 39.1%; Score 127; DB 1; Length 67;

Best Local Similarity 54.7%; Pred. No. 4.4e-08;

Matches 29; Conservative 4; Mismatches 20; Indels 0; Gaps 0;

QY 1 MRCPLVLIILLLTASAPGVVLPKTEDDVPMSVYGNKSGILRLNGVCC 53

||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 MRCFPVFIILLIIASAPCFDARTKDDDDVPLSLRDLNKRITRLNRECC 53

RESULT 7

CX51_CONIM

ID CX51_CONIM STANDARD; PRT; 62 AA.

AC Q9U6Z5;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Putative conotoxin IM5.1 precursor.

OS *Conus imperialis* (Imperial cone).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;

OC Neogastropoda; Conoidea; Conidae; Conus.

OX NCBI_TaxID=35631;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Venom duct;

RX MEDLINE=99452958; PubMed=10521453;

RA Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,

RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,

RA Bandyopadhyay P., Craig A.G., Olivera B.M.;

RT "The T-superfamily of conotoxins.";

RL J. Biol. Chem. 274:30664-30671(1999).

RN [2]

RP ERRATUM.

RA Walker C.S., Steel D., Jacobsen R.B., Lirazan M.B., Cruz L.J.,

RA Hooper D., Shetty R., Delacruz R.C., Nielsen J.S., Zhou L.M.,

RA Bandyopadhyay P., Craig A.G., Olivera B.M.;

RL J. Biol. Chem. 274:36030-36030(1999).

CC -!- FUNCTION: Not known. Has only been studied at nucleotidic level.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.

CC -!- SIMILARITY: BELONGS TO THE T-SUPERFAMILY OF CONOTOXINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF167169; AAF03689.1; -.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 45

FT PEPTIDE 46 60 PUTATIVE CONOTOXIN IM5.1.

FT DISULFID 50 57 POTENTIAL.

FT DISULFID 51 58 POTENTIAL.

FT MOD_RES 60 60 AMIDATION (G-61 PROVIDE AMIDE GROUP)

FT (POTENTIAL).

SQ SEQUENCE 62 AA; 7030 MW; 9570E7C02275796D CRC64;

Query Match 32.0%; Score 104; DB 1; Length 62;

Best Local Similarity 46.3%; Pred. No. 1.8e-05;

Matches 25; Conservative 7; Mismatches 20; Indels 2; Gaps 1;

QY 1 MRCPLVLIILLLTASAPGVVLPKTEDDVPMSVYGNKSGILRLNGVCC 54

||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1 MYCLPVFIILLIISSAPSTPPQPRNKDRVHLISLLDNHKKIOLRDWNS--CCG 52

RESULT 8

CX44_CONMR

ID CX44_CONMR STANDARD; PRT; 13 AA.

AC P58810;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lambda/chi-conotoxin MrIB (Chi-MrIB).

OS *Conus marmoreus* (Marble cone).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;

OC Neogastropoda; Conoidea; Conidae; Conus.

OX NCBI_TaxID=42752;

RN [1]

RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.

RC TISSUE=Venom;

```

RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Pallant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- PTM: Exists in two forms, due to cis-trans isomerization at His-
CC 11-Hyp-12.
CC -!- MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB; 1IEO; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT DISULFID 5 10
FT MOD_RES 12 12
SQ SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 25.5%; Score 83; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 GVCCGYKLCHPC 61
DB 2 GVCCGYKLCHPC 13

RESULT 9
CXLL_CONMR STANDARD; PRT; 11 AA.
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CmrVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT DISULFID 3 8
FT MOD_RES 10 10
SQ SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;

Query Match 23.7%; Score 77; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 VCCGYKLCHPC 61
DB 1 VCCGYKLCHPC 11

RESULT 10

```

```

BV8_BOMVA STANDARD; PRT; 96 AA.
ID BV8_BOMVA
AC Q9PW66;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein Bv8 precursor.
OS Bombina variegata (Yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=99349621; PubMed=10422759;
RA Mollay C., Wechsberger C., Mignogna G., Negri L., Melchiorri P.,
RA Barra D., Kreil G.;
RT "Bv8, a small protein from frog skin and its homologue from snake
RT venom induce hyperalgesia in rats.";
RL Eur. J. Pharmacol. 374:189-196(1999).
CC -!- FUNCTION: Potentially contract gastrointestinal (GI) smooth muscle.
CC Induces hyperalgesia.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE PROKINECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF168790; AAD45816.1;
DR HSP; P25687; IIMT.
KW SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 96
FT DISULFID 26 38
FT DISULFID 32 50
FT DISULFID 37 78
FT DISULFID 60 86
FT DISULFID 80 95
SQ SEQUENCE 96 AA; 10102 MW; A12490A7437609B4 CRC64;

Query Match 19.7%; Score 64; DB 1; Length 96;
Best Local Similarity 22.1%; Pred. No. 1.1;
Matches 17; Conservative 8; Mismatches 20; Indels 32; Gaps 2;

QY 1 MRCPLVLIILLITASAPGVVLPKTEDDVPMSVYGVNGKSLRILRNGVCCGYKL--- 57
DB 1 MKCFAQIVVLLLVIAFSGHGVITGACDKDVQCGS-----GTCCAASAWSR 45
QY 58 -----CHP 60
DB 46 NIRFCIPLNGSGEDCHP 62

RESULT 11
MCTL_CANFA STANDARD; PRT; 249 AA.
ID MCTL_CANFA
AC P21842;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chymase precursor (EC 3.4.21.39) (Mast cell protease I).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90335214; PubMed=2378872;

```


DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Properdin precursor (Factor P).
GN PFC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93038568; PubMed=1417780;
RA Nolan K.F., Kaluz S., Higgins J.M., Goundis D., Reid K.B.M.;
RT "Characterization of the human properdin gene.";
RL Biochem. J. 287:291-297(1992).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=91184288; PubMed=2009915;
RA Nolan K.F., Schwaible W., Kaluz S., Dierich M.P., Reid K.B.M.;
RT "Molecular cloning of the cDNA coding for properdin, a positive
regulator of the alternative pathway of human complement.";
RL Eur. J. Immunol. 21:771-776(1991).
RN [3]
RN SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A., AND VARIANT PFC TYPE-II DEFICIENCY TRP-100.
RX MEDLINE=96079085; PubMed=8530058;
RA Westberg J., Fredrikson G.N., Truedsson L., Sjoeholm A.G., Uhlen M.;
RT "Sequence-based analysis of properdin deficiency: identification of
point mutations in two phenotypic forms of an X-linked
immunodeficiency.";
RL Genomics 29:1-8(1995).
RN [5]
RN SEQUENCE FROM N.A.
RP TISSUE-Spleen;
RC Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RN CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=20435812; PubMed=10878002;
RA Hartmann S., Hofsteenge J.;
RT "Properdin, the positive regulator of complement, is highly
C-mannosylated.";
RL J. Biol. Chem. 275:28569-28574(2000).
RN [7]
RN VARIANT PFC TYPE-III DEFICIENCY ASP-414.
RX MEDLINE=97025478; PubMed=8871668;
RA Fredrikson G.N., Westberg J., Kuijper E.J., Tijssen C.C.,
RA Sjoeholm A.G., Uhlen M., Truedsson L.;
RT "Molecular characterization of properdin deficiency type III:
dysfunction produced by a single point mutation in exon 9 of the
structural gene causing a tyrosine to aspartic acid interchange.";
RL J. Immunol. 157:3666-3671(1996).
RN [8]
RN VARIANT PFC TYPE-II DEFICIENCY ARG-343.
RX MEDLINE=96376603; PubMed=9710744;
RA Fredrikson G.N., Gullstrand B., Westberg J., Sjoeholm A.G., Uhlen M.,
RA Truedsson L.;
RT "Expression of properdin in complete and incomplete deficiency: normal
in vitro synthesis by monocytes in two cases with properdin
deficiency type II due to distinct mutations.";
RL J. Clin. Immunol. 18:272-282(1998).
RN [9]
RN VARIANT PFC TYPE-I DEFICIENCY VAL-298.
RX MEDLINE=20365201; PubMed=10909851;
RA van den Bogaard R., Fijen C.A.P., Schipper M.G.J., de Galan L.,
RA Kuijper E.J., Mannens M.M.A.M.;
RT "Molecular characterisation of 10 Dutch properdin type I deficient
families: mutation analysis and X-inactivation studies.";
RL Eur. J. Hum. Genet. 8:513-518(2000).
CC -!- FUNCTION: A POSITIVE REGULATOR OF THE ALTERNATE PATHWAY OF
COMPLEMENT. IT BINDS TO AND STABILIZES THE C3-AND C5-CONVERTASE

CC ENZYME COMPLEXES.
CC -!- DISEASE: DEFICIENCY OF PROPERDIN LEADS TO HIGHER SUSCEPTIBILITY TO
BACTERIAL INFECTIONS; ESPECIALLY TO MENINGOCOCCAL INFECTIONS.
CC THREE PHENOTYPES HAVE BEEN REPORTED: COMPLETE DEFICIENCY (TYPE I),
CC INCOMPLETE DEFICIENCY (TYPE II), AND DYSFUNCTION OF PROPERDIN
CC (TYPE III).
CC -!- SIMILARITY: CONTAINS 6 TSP TYPE-1 DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83652; AAA36489.1; -
CC EMBL; X70872; CAA50220.1; -
CC EMBL; X57748; CAA40914.1; -
CC EMBL; AL009172; CAA15658.1; ALT_SEQ.
CC EMBL; AF005665; AAB63280.1; -
CC EMBL; AF005666; AAC51626.1; -
CC EMBL; AF005668; AAB62886.1; -
CC EMBL; AF005667; AAB62886.1; JOINED.
CC EMBL; AF005664; AAB63279.1; -
CC EMBL; BC015756; AAB15756.1; -
CC PIR; S16150; S16150.
CC PIR; S29126; S29126.
CC Genew; HGNC:8864; PFC.
CC MIM; 300383; -
CC MIM; 312060; -
CC InterPro; IPR000884; TSP1.
CC Pfam; PF00090; tsp_1; 5.
CC SMART; SM00209; TSP1; 6.
CC PROSITE; PS50092; TSP1; 6.
KW Signal; Complement alternate pathway; Glycoprotein; Repeat;
KW Disease mutation.
FT SIGNAL 1 27
FT CHAIN 28 469
FT DOMAIN 77 134
FT DOMAIN 135 191
FT DOMAIN 192 255
FT DOMAIN 256 313
FT DOMAIN 314 377
FT DOMAIN 378 437
FT CARBOHYD 83 83
FT CARBOHYD 86 86
FT CARBOHYD 139 139
FT CARBOHYD 142 142
FT CARBOHYD 145 145
FT CARBOHYD 196 196
FT CARBOHYD 199 199
FT CARBOHYD 260 260
FT CARBOHYD 263 263
FT CARBOHYD 321 321
FT CARBOHYD 324 324
FT CARBOHYD 382 382
FT CARBOHYD 385 385
FT CARBOHYD 388 388
FT CARBOHYD 428 428
FT CARBOHYD 457 457
FT VARIANT 100 100
FT VARIANT 298 298
FT VARIANT 343 343
FT VARIANT 414 414
FT CONFLICT 457 457
FT SEQUENCE 469 AA; 51276 MW; 5EB42B63F0283917 CRC64;
Query Match 18.6%; Score 60.5; DB 1; Length 469;
Best Local Similarity 29.2%; Pred. No. 13;


```
OX NCBI_TaxID=562, 83334;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN-K12;
RA MEDLINE=92283811; PubMed=1317854;
RA Ta D.T., Vickery L.E.;
RT "Cloning, sequencing, and overexpression of a [2Fe-2S] ferredoxin
RL J. Biol. Chem. 267:11120-11125(1992).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA MEDLINE=94131939; PubMed=8300516;
RA Kawula T.H., Lelivelt M.J.;
RT "Mutations in a gene encoding a new Hsp70 suppress rapid DNA
RT inversion and bgl activation, but not proU derepression, in hns-1
RT mutant Escherichia coli.";
RL J. Bacteriol. 176:610-619(1994).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RA MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampel G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horluchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN (5)
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimallanta E.T., Potamousis K.,
RA Apodaca J., Anantharam T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN (6)
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RA MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN (7)
RP SEQUENCE OF 1-7.
RA MEDLINE=75112169; PubMed=4375562;
RA Kneoll H.-E., Knappe J.;
RT "Escherichia coli ferredoxin, an iron-sulfur protein of the
RT adrenodoxin type.";
RL Eur. J. Biochem. 50:245-252(1974).
CC -!- FUNCTION: FERREDOXIN ARE IRON-SULFUR PROTEINS THAT TRANSFER
```

```
CC CC
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS. ALTHOUGH THE
CC FUNCTION OF THIS FERREDOXIN IS UNKNOWN IT IS PROBABLE THAT IT
CC HAS A ROLE AS A CELLULAR ELECTRON TRANSFER PROTEIN.
CC -!- COFACTOR: BINDS 1 2FE-2S CLUSTER.
CC -!- SIMILARITY: BELONGS TO THE ADRENODOXIN / PUTIDAREDOXIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M88654; AAA23755.1; -.
CC EMBL; U05338; AAD13474.1; -.
CC EMBL; U01827; -. NOT_ANNOTATED_CDS.
CC EMBL; AE000339; AAC75578.1; -.
CC EMBL; D90882; BAA16415.1; -.
CC EMBL; D90883; BAA16419.1; -.
CC EMBL; AE005482; AAG57639.1; -.
CC EMBL; AP002562; BAB36814.1; -.
CC PIR; JCL110; JC1110.
CC HSPG; P00259; IGPX.
CC ECGene; EGL1328; fdx.
CC InterPro; IPR001055; Adrenodoxin.
CC InterPro; IPR001041; Ferredoxin.
CC Pfam; PF00111; fer2.1
CC PRINTS; PR00355; ADRENODOXIN.
CC PROSITE; PS00814; ADX; 1.
CC Electron transport; Iron-sulfur; Complete proteome.
CC INIT_MET 0
CC METAL 41 41 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 47 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 50 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 86 86 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC CONFLICT 7 7 P -> Y (IN REF. 7).
CC CONFLICT 62 62 P -> Q (IN REF. 2).
CC CONFLICT 67 68 QE -> HQ (IN REF. 2).
CC CONFLICT 73 73 D -> N (IN REF. 2).
CC CONFLICT 76 76 W -> R (IN REF. 2).
CC CONFLICT 90 92 VTD -> SYH (IN REF. 2).
CC SEQUENCE 110 AA; 12199 MW; 6CDE92D4819FAA75 CRC64;
Query Match 18.3%; Score 59.5; DB 1; Length 110;
Best Local Similarity 32.7%; Pred No. 4.3;
Matches 17; Conservative 8; Mismatches 16; Indels 11; Gaps 3;
QY 18 PGVVVLPTKTEDVPMSSVY--GNGKSILRGILRNGV-----CCGYKLCH 59
Db 1 PRIVILPH-QDLCPDGAVLEANSGETILDALRNGIEIEHACEKSCACTTCH 51
Search completed: January 29, 2003, 09:29:25
Job time : 15.8636 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:57 ; Search time 25.6477 Seconds
(without alignments)
228.644 Million cell updates/sec

Title: US-09-580-201a-12
Perfect score: 325
Sequence: 1 MRCLPVLIIILLITASAPGV.....ILRGILRNGVCCGYKLCHPC 61
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	148.5	45.7	62	D59147	conotoxin Gm5.2 pr
2	140	43.1	63	C59147	conotoxin Gm5.1 pr
3	138	42.5	62	B59147	conotoxin Tx5.1 pr
4	136	41.8	63	F59147	conotoxin P5.1 pr
5	127	39.1	67	E59147	conotoxin Tx5.2 pr
6	104	32.0	62	G59147	conotoxin Im5.1 pr
7	65.5	20.2	350	E83236	hypothetical prote
8	64	19.7	453	T20965	hypothetical prote
9	64	19.7	477	T47553	hypothetical prote
10	63	19.4	126	T42321	hypothetical prote
11	61	18.8	249	A35842	chymase (EC 3.4.21
12	60.5	18.6	112	B83168	ferredoxin [2Fe-2S
13	60.5	18.6	469	S29126	properdin precursor
14	60	18.5	258	T30368	probable immediate
15	60	18.5	570	A63458	urase (EC 3.5.1.5
16	60	18.5	729	T51896	probable translati
17	60	18.5	1324	T01508	mismatch repair en
18	59.5	18.3	111	JC1110	ferredoxin [2Fe-2S
19	59.5	18.3	111	C85897	[2FE-2S] ferredoxi
20	59.5	18.3	111	G10552	[2FE-2S] ferredoxi
21	59.5	18.3	114	S77591	ferredoxin [2Fe-2S
22	59	18.2	279	T37061	probable secreted
23	59	18.2	379	S49614	cych protein - Rhi
24	59	18.2	1322	T15689	hypothetical prote
25	58.5	18.0	221	F71985	carbonic anhydrase
26	58.5	18.0	262	S38620	allergen Phl p I -
27	58.5	18.0	263	S44182	conserved hypothet
28	58.5	18.0	311	H75411	hypothetical prote
29	58.5	18.0	470	T49272	

30	58.5	18.0	913	2	AC2445	hypothetical prote
31	58.5	18.0	2970	2	T08839	polyprotein - marm
32	58	17.8	79	2	F95210	conserved domain p
33	58	17.8	374	2	T05571	hypothetical prote
34	57.5	17.7	106	1	IPXL2	insulin II precurs
35	57.5	17.7	775	2	T00962	hypothetical prote
36	57	17.5	325	2	B43692	T2 protein - rabbi
37	57	17.5	405	2	B69200	hypothetical prote
38	57	17.5	466	2	A37351	IMP-E2 protein pre
39	57	17.5	508	2	D81325	probable phosphate
40	57	17.5	684	2	B69308	conserved hypothet
41	57	17.5	1353	1	J02168	E2 glycoprotein pr
42	56.5	17.4	504	2	F83603	threonine dehydrat
43	56.5	17.4	521	2	T27192	hypothetical prote
44	56.5	17.4	850	2	C83081	probable oxidoredu
45	56	17.2	239	2	AH0541	probable outer mem

ALIGNMENTS

RESULT 1

D59147
conotoxin Gm5.2 precursor - cone shell (Conus gloriamaris)
C:Species: Conus gloriamaris (glory-of-the-sea cone)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
C:Accession: D59147
R:Walker, C.S.; Steel, D.; Jacobsen, R.B.; Lirazan, M.B.; Cruz, L.J.; Hooper, D.; She
J. Biol. Chem. 274, 30664-30671, 1999
A>Title: The T-superfamily of conotoxins.
A:Reference number: A59147; MUID:99452958; PMID:10521453
A:Accession: D59147
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-62 <WAL>
A:Cross-references: GB:AF167166; NID:g6103610; PIDN:AAF03686.1; PID:g6103611
A>Note: submitted to GenBank, July 1999
C:Superfamily: unassigned conotoxins
C:Keywords: amidated carboxyl end; toxin; venom
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-49/Domain: propeptide #status predicted <PRO>
F:50-60/Product: conotoxin gm5 #status predicted <MAT>
F:51-58,52-59/Disulfide bonds: #status predicted
F:60/Modified site: amidated carboxyl end (Ser) (amide in mature form from following

Query Match	45.7%	Score	148.5	DB	2	Length	62
Best Local Similarity	58.5%	Pred. No.	3.4e-10				
Matches	31	Conservative	8	Mismatches	13	Indels	1
Gaps	1						
QY	1	MRCLPVLIIILLITASAPGVVLPKTEDDVPMSSVYGNKSLIRGLRNGVCC	53				
Db	1	MRCLPVEVILLIIILASAPSDAQPDKDDVPLPLHDNIRSTLQ-TLRKKVCC	52				

RESULT 2

C59147
conotoxin Gm5.1 precursor - cone shell (Conus gloriamaris)
C:Species: Conus gloriamaris (glory-of-the-sea cone)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
C:Accession: C59147
R:Walker, C.S.; Steel, D.; Jacobsen, R.B.; Lirazan, M.B.; Cruz, L.J.; Hooper, D.; She
J. Biol. Chem. 274, 30664-30671, 1999
A>Title: The T-superfamily of conotoxins.
A:Reference number: A59147; MUID:99452958; PMID:10521453
A:Accession: C59147
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-63 <WAL>
A:Cross-references: GB:AF167165; NID:g6103608; PIDN:AAF03685.1; PID:g6103609
A>Note: submitted to GenBank, July 1999
C:Superfamily: unassigned conotoxins
C:Keywords: toxin; venom
F:1-19/Domain: signal sequence #status predicted <SIG>

F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-45/Domain: propeptide #status predicted <PRO>
F:46-60/Product: conotoxin im5 #status predicted <MAT>
F:50-57, 51-58/disulfide bonds: #status predicted
F:60/Modified site: amidated carboxyl end (Trp) (amide in mature form from following gly

Query Match 32.0%; Score 104; DB 2; Length 62;
Best Local Similarity 46.3%; Pred. No. 4.7e-05;
Matches 25; Conservative 7; Mismatches 20; Indels 2; Gaps 1;

QY 1 MRCLPVLIIILLTASAPGVVLPKTEDDVPMSSVYGNGKSTILRLNGVCCG 54
Db 1 MYCLPFIILLIISASFPPTPPQRNKRDKVHLISLDNHKQILQRDWSN--CCG 52

RESULT 7
E83236
hypothetical protein PA3286 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83236
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <STO>
A:Cross-references: GB:AE004750; GB:AE004091; NID:g9949400; PIDN:AAG06674.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3286

Query Match 20.2%; Score 65.5; DB 2; Length 350;
Best Local Similarity 36.7%; Pred. No. 6.6;
Matches 18; Conservative 9; Mismatches 17; Indels 5; Gaps 1;

QY 5 PVLIIILLTASAPGVVLPKTEDDVPMSSV-----YNGKSTILRLGILR 48
Db 300 PVLDSYANTSSAGSVIALHKKHDDLPSCAIGVLSFGAGYSIGSVILR 348

RESULT 8
T20965
hypothetical protein F15B9.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T20965
R:Percy, C.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19351
A:Accession: T20965
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-453 <WIL>
A:Cross-references: EMBL:Z78013; PIDN:CAB01424.1; GSPDB:GN00023; CESP:F15B9.6
A:Experimental source: clone F15B9
C:Genetics:
A:Gene: CESP:F15B9.6
A:Map position: 5
A:Introns: 17/1; 74/3; 138/3; 209/1; 240/3; 275/1; 335/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F15B9.6

Query Match 19.7%; Score 64; DB 2; Length 453;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 19; Conservative 11; Mismatches 22; Indels 24; Gaps 3;

QY 1 MRCLPVLIIILLTASAPGVV-----LPKTEDDVPMSSVYGNGKSTILRG 45
Db 1 MKCLLYFIVILIQANVFGSDSPAYNILEPNLGGCPRTSDRIKFCSAKFNS----- 55

QY 46 ILRNGV---CCGYKLC 58
Db 56 -IRNKVSPITCAHQC 70

RESULT 9
T47753
hypothetical protein F24I3.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47753
R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24475
A:Accession: T47753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <NYA>
A:Cross-references: EMBL:AL138655
A:Experimental source: cultivar Columbia; BAC clone F24I3
C:Genetics:
A:Map position: 3
A:Introns: 19/3; 82/3; 185/3; 263/3
A:Note: F24I3.10

Query Match 19.7%; Score 64; DB 2; Length 477;
Best Local Similarity 28.9%; Pred. No. 13;
Matches 24; Conservative 9; Mismatches 26; Indels 24; Gaps 5;

QY 3 CLPVL-----IILLLTASA-PGVV---VLPKTEDDVP-----MSSVYGNGKSI 42
Db 69 CIPILCVSWILFDIFFLLMTSSRDGIVPRSFRRPPTDDADPTSPMSMEWVGRTPTNI 128

QY 43 ----LRGILRNGVCCGYKLC 61
Db 129 RIPRVKDVTVNGHTVVKRFDTC 151

RESULT 10
T42321
hypothetical protein 31.1 - phage SPPI
C:Species: phage SPPI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: T42321
R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.
Gene 204, 201-212, 1997
A:Title: The complete nucleotide sequence and functional organization of Bacillus sub
A:Reference number: Z22137; MUID:98094274; PMID:9434185
A:Accession: T42321
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-126 <ALO>
A:Cross-references: EMBL:X97918; PIDN:CAA66528.1
C:Superfamily: phage SPPI hypothetical protein 31.1

Query Match 19.4%; Score 63; DB 2; Length 126;
Best Local Similarity 24.0%; Pred. No. 4.9;
Matches 18; Conservative 9; Mismatches 14; Indels 34; Gaps 2;

QY 13 LTASAPGVVLPKTEDDVPMSSVYGNGKSTILRGILRN----- 49
Db 42 IKASLPFLILPKSALAAGIDSTFGN----VHGAINMADAGVVLVIFAGAAGLGNRT 97

QY 50 -----GVCCGYKL 57
Db 98 QAIEILIGVCCGYIL 112

RESULT 11
A35842
chymase (EC 3.4.21.39) precursor - dog
N:Alternate names: mast cell protease I; skeletal muscle (sk) protease

Search completed: January 29, 2003, 09:33:01

THIS PAGE BLANK (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 61 Seconds
(without alignments)
133.251 Million cell updates/sec

Title: US-09-580-201A-12
Perfect score: 325
Sequence: 1 MRCLPVLIIILLTASAPGV.....ILRGIIRNGVCCGYKLCHPC 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	325	100.0		61	21	AA0808016	Amino acid sequence
2	316	97.2		61	21	AA092231	Chi-conotoxin, chi
3	154	47.4		59	21	AA097123	Tau conotoxin prot
4	153	47.1		60	21	AA097124	Tau conotoxin prot
5	153	47.1		70	21	AA097122	Tau conotoxin prot
6	148.5	45.7		62	21	AA097129	Tau conotoxin prot
7	146.5	45.1		63	21	AA097126	Tau conotoxin prot
8	140	43.1		61	21	AA097116	Tau conotoxin prot
9	140	43.1		61	21	AA097119	Tau conotoxin prot
10	140	43.1		63	21	AA097128	Tau conotoxin prot

11	138	42.5	62	21	AA097115	Tau conotoxin prot
12	136.5	42.0	64	21	AA097125	Tau conotoxin prot
13	130	40.0	67	21	AA097121	Tau conotoxin prot
14	128	39.4	62	21	AA097127	Tau conotoxin prot
15	127	39.1	67	21	AA097120	Tau conotoxin prot
16	125	38.5	62	21	AA097117	Tau conotoxin prot
17	104	32.0	62	21	AA097118	Tau conotoxin prot
18	89	27.4	13	21	AA08017	Amino acid sequenc
19	89	27.4	13	21	AA08017	Chi-conotoxin pept
20	83	25.5	12	21	AA08018	Amino acid sequenc
21	83	25.5	13	21	AA092230	Chi-conotoxin pept
22	77	23.7	12	21	AA08019	Amino acid sequenc
23	75	23.1	13	21	AA08013	Generic formula fo
24	69	21.2	12	21	AA08014	Generic formula fo
25	69	21.2	12	21	AA08015	Generic formula fo
26	64	19.7	174	11	AA07664	Cysteine-added var
27	63.5	19.5	625	22	AA093033	C glutamicum prote
28	63	19.4	75	21	AA090575	Conus consors kapp
29	63	19.4	75	21	AA090586	Conus sulcatus kap
30	63	19.4	75	21	AA090587	Conus sulcatus kap
31	62	19.1	69	18	AA012764	A-lineage conotoxi
32	61.5	18.9	38	22	AB038490	Peptide #5996 enco
33	61.5	18.9	38	22	AB023643	Protein #5642 enco
34	61.5	18.9	38	22	AA053105	Human brain expres
35	61.5	18.9	38	22	AA071640	Human bone marrow
36	61.5	18.9	38	22	AA019246	Peptide #5680 enco
37	61.5	18.9	38	22	AA031934	Peptide #5971 enco
38	61.5	18.9	38	23	AB041454	Human peptide enco
39	61	18.8	64	21	AA090583	Conus striatus kap
40	61	18.8	96	23	AA021287	Corn ACP-like prot
41	61	18.8	161	23	AB065215	Human albumin fusi
42	61	18.8	161	23	AB065216	Human albumin fusi
43	61	18.8	161	23	AA091095	Human secreted pro
44	61	18.8	161	23	AA091125	Human secreted pro
45	61	18.8	317	13	AA023600	Recombinant hemato

ALIGNMENTS

RESULT 1
AA08016
ID AAB08016 standard; Protein; 61 AA.
XX
AC AAB08016;
XX
DT 14-NOV-2000 (first entry)
XX
DE Amino acid sequence of the conotoxin Marl propeptide.
XX
KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
XX
OS Conus marmoreus.
XX
PN WO200044769-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US01978.
XX
PR 29-JAN-1999; 99US-0118381.
PR 28-DEC-1999; 99US-0173343.
XX
(UTAH) UNIV UTAH RES FOUND.
PA
PI McIntosh JM, Olivera BM, Cruz LJ;
XX
DR WPI; 2000-476222/41.
XX
N-PSDB; AAA63513.
PT Purified ap-conotoxin derived from cone snail venom for use as an
analgesic -
XX

PS Claim 23; Page 13-14; 29pp; English.

XX The present sequence represents a Marl propeptide. Marl is an
 CC ap-conotoxin peptide. Conotoxins are naturally available in minute
 CC amounts in the venom of cone snails. The peptides have analgesic
 CC activity. The peptides are used to treat or prevent pain.

XX
 SQ Sequence 61 AA;

Query Match 100.0%; Score 325; DB 21; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.7e-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSSVYGNKSLILGRNGVCCGYKLCHP 60
 |||||
 Db 1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSSVYGNKSLILGRNGVCCGYKLCHP 60

QY 61 C 61
 |
 Db 61 C 61

RESULT 2

AAY92231
 ID AAY92231 standard; Protein: 61 AA.

XX
 AC AAY92231;

XX
 DT 10-AUG-2000 (first entry)

DE Chi-conotoxin, chi-MrIA, leader and mature peptide.

XX
 KW chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenaline transporter; urinary tract; analgesic; cardiant;
 KW antiarrhythmic; antidepressant; anxiolytic; anti-inflammatory.

XX
 OS Conus marmoreus.

XX
 PN WO200020444-A1.

XX
 PD 13-APR-2000.

XX
 PF 01-OCT-1999; 99WO-AU00844.

XX
 PR 02-OCT-1998; 98AU-0006274.

XX
 PA (UYQU) UNIV QUEENSLAND.

XX
 PI Lewis RJ, Alewood PF, Sharpe IA;

XX
 WPI: 2000-303738/26.

DR
 DR N-PSDB; ARA09112.

XX
 PT Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation

PS Example 7; Page 31; 47pp; English.

XX This sequence is the conotoxin, chi-MrIA, a member of a new class of
 CC conotoxins, designated chi-conotoxin. It was isolated from the venom of
 CC the mollusc hunting cone snail, Conus marmoreus. The peptide is an
 CC inhibitor of the neuronal amine transporters, especially the neuronal
 CC noradrenaline transporter. Inhibitors of noradrenaline re-uptake which
 CC have a negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chi-MrIA (0.1 nmol micro M)
 CC inhibited the accumulation of radiolabeled noradrenaline in a
 CC concentration-dependent manner, with a log IC₅₀ value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are

CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).

XX
 SQ Sequence 61 AA;

Query Match 97.2%; Score 316; DB 21; Length 61;
 Best Local Similarity 98.4%; Pred. No. 1.8e-29;
 Matches 60; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSSVYGNKSLILGRNGVCCGYKLCHP 60
 |||||
 Db 1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSSVYGNKSLILGRNGVCCGYKLCHP 60

QY 61 C 61
 |
 Db 61 C 61

RESULT 3

AAY97123

ID AAY97123 standard; Protein: 59 AA.

XX
 AC AAY97123;

XX
 DT 22-DEC-2000 (first entry)

DE Tau conotoxin protein precursor.

XX
 KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
 KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
 KW migraine; treatment; therapy.

XX
 OS Conus marmoreus.

XX
 PN WO200046371-A1.

XX
 PD 10-AUG-2000.

XX
 PF 04-FEB-2000; 2000WO-US03021.

XX
 PR 04-FEB-1999; 99US-0118642.

XX
 PA (UTAH) UNIV UTAH RES FOUND.

XX
 PA (COGN-) COGNETIX INC.

XX
 PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
 PI Jones RM;

XX
 WPI: 2000-543489/49.

DR
 DR N-PSDB; AAA51980.

XX Novel analgesic short peptides from predatory cone snails for treating
 PT acute, chronic and neuropathic pain and migraines

PS Claim 33; Page 20; 58pp; English.

XX The venom of predatory cone snails comprises relatively small
 CC peptides which are targeted to various neuromuscular receptors
 CC and may be equivalent in their pharmacological diversity to the
 CC alkaloids of plants or secondary metabolites of microorganisms.
 CC Several peptides have been characterised from Conus venoms. These
 CC include the alpha and mu conotoxins which target nicotinic
 CC acetylcholine receptors and muscle sodium channels respectively.
 CC Chronic or intractable pain as well as neuropathic pain is
 CC currently treated with a range of analgesic compounds. It is
 CC thought that the tau-conotoxins described could have applications
 CC as analgesic drugs and could be used for treating such acute,
 CC chronic and neuropathic pain as well as migraines.

XX
 SQ Sequence 59 AA;

AA97122	
ID	AA97122 standard; Protein; 70 AA.
XX	XX
AC	AA97122;
XX	XX
DT	22-DEC-2000 (first entry)
XX	XX
DE	Tau conotoxin protein precursor.
XX	XX
KW	Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
KW	antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
KW	migraine; treatment; therapy.
XX	XX
OS	Conus marmoreus.
XX	XX
PN	WC2000046371-A1.
XX	XX
PD	10-AUG-2000.
XX	XX
PF	04-FEB-2000; 2000WO-US03021.
XX	XX
PR	04-FEB-1999; 99US-0118642.
XX	XX
PA	(UTAH) UNIV UTAH RES FOUND.
PA	(COGN-) COGNETIX INC.
XX	XX
PI	Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
PI	Jones RM;
XX	XX
DR	WPI; 2000-543489/49.
XX	XX
DR	N-PSDB; AAA51979.
XX	XX
PT	Novel analgesic short peptides from predatory cone snails for treating
PT	acute, chronic and neuropathic pain and migraines
XX	XX
PS	Claim 33; Page 20; 58pp; English.
XX	XX
CC	The venom of predatory cone snails comprises relatively small
CC	peptides which are targeted to various neuromuscular receptors
CC	and may be equivalent in their pharmacological diversity to the
CC	alkaloids of plants or secondary metabolites of microorganisms
CC	Several peptides have been characterised from Conus venoms. These
CC	include the alpha and mu conotoxins which target nicotinic
CC	acetylcholine receptors and muscle sodium channels respectively.
CC	Chronic or intractable pain as well as neuropathic pain is
CC	currently treated with a range of analgesic compounds. It is
CC	thought that the tau-conotoxins described could have applications
CC	as analgesic drugs and could be used for treating such acute,
CC	chronic and neuropathic pain as well as migraines.
XX	XX
SQ	Sequence 70 AA;
Query Match	47.1%; Score 153; DB 21; Length 70;
Best Local Similarity	54.2%; Pred. NO. 2.1e-10;
Matches 32; Conservative	7; Mismatches 14; Indels 6; Gaps
QY	1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSVYVNGKSLR-----GILRNGVCC 53
	: : : :
	: : : :
Db	1 MRCLPVLIIILLITASAPSDARPKTKDMPPLASFDNAKRILQILODRNCCRGDCC 59
RESULT 6	
AA97129	
ID	AA97129 standard; Protein; 62 AA.
XX	XX
AC	AA97129;
XX	XX
DT	22-DEC-2000 (first entry)
XX	XX
DE	Tau conotoxin protein precursor.
XX	XX
KW	Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
KW	antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;

DR WPI: 2000-543489/49.
 DR N-PSDB; AAA51973.
 PT Novel analgesic short peptides from predatory cone snails for treating
 PT acute, chronic and neuropathic pain and migraines
 XX
 PS Claim 33; Page 17-18; 58pp; English.
 CC The venom of predatory cone snails comprises relatively small
 CC peptides which are targeted to various neuromuscular receptors
 CC and may be equivalent in their pharmacological diversity to the
 CC alkaloids of plants or secondary metabolites of microorganisms.
 CC Several peptides have been characterised from Conus venoms. These
 CC include the alpha and mu conotoxins which target nicotinic
 CC acetylcholine receptors and muscle sodium channels respectively.
 CC Chronic or intractable pain as well as neuropathic pain is
 CC currently treated with a range of analgesic compounds. It is
 CC thought that the tau-conotoxins described could have applications
 CC as analgesic drugs and could be used for treating such acute,
 CC chronic and neuropathic pain as well as migraines.
 XX
 SQ Sequence 61 AA;
 Query Match 43.1%; Score 140; DB 21; Length 61;
 Best Local Similarity 56.4%; Pred. No. 5.8e-09;
 Matches 31; Conservative 6; Mismatches 14; Indels 4; Gaps 2;
 QY 1 MRLCPVLIIILLITASAPGVVLPKTEDDVPMSVYG--NGKSILRGILRNGVCC 53
 | ||||:||||| ||| | ||||| ||||:| | : ||: | |||
 Db 1 MCCLPVFVILLITLTSAPSDALPKTRDDVPLASPHGGYNARRILQ--RRQGWCC 53
 ||||:||||| ||| | ||||| ||||:| | : ||: | |||
 RESULT 10
 AAY97128
 ID AAY97128 standard; Protein; 63 AA.
 XX
 AC AAY97128;
 XX
 DT 22-DEC-2000 (first entry)
 XX
 DE Tau conotoxin protein precursor.
 XX
 KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
 KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
 KW migraine; treatment; therapy.
 XX
 OS Conus gloriamaris.
 XX
 PN WO200046371-A1.
 XX
 PD 10-AUG-2000.
 XX
 PF 04-FEB-2000; 2000WO-US03021.
 XX
 PR 04-FEB-1999; 99US-0118642.
 XX
 PA (UTAH) UNIV UTAH RES FOUND.
 PA (COGN-) COGNETIX INC.
 XX
 PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
 PI Jones RM;
 XX
 DR WPI: 2000-543489/49.
 DR N-PSDB; AAA51985.
 XX
 PT Novel analgesic short peptides from predatory cone snails for treating
 PT acute, chronic and neuropathic pain and migraines
 XX
 PS Claim 33; Page 22-23; 58pp; English.
 XX
 CC The venom of predatory cone snails comprises relatively small
 CC peptides which are targeted to various neuromuscular receptors
 CC and may be equivalent in their pharmacological diversity to the
 CC alkaloids of plants or secondary metabolites of microorganisms.
 CC Several peptides have been characterised from Conus venoms. These
 CC include the alpha and mu conotoxins which target nicotinic
 CC acetylcholine receptors and muscle sodium channels respectively.
 CC Chronic or intractable pain as well as neuropathic pain is
 CC currently treated with a range of analgesic compounds. It is
 CC thought that the tau-conotoxins described could have applications
 CC as analgesic drugs and could be used for treating such acute,
 CC chronic and neuropathic pain as well as migraines.
 XX
 SQ Sequence 63 AA;
 Query Match 43.1%; Score 140; DB 21; Length 61;
 Best Local Similarity 56.4%; Pred. No. 5.8e-09;
 Matches 31; Conservative 6; Mismatches 14; Indels 4; Gaps 2;
 QY 1 MRLCPVLIIILLITASAPGVVLPKTEDDVPMSVYG--NGKSILRGILRNGVCC 53
 | ||||:||||| ||| | ||||| ||||:| | : ||: | |||
 Db 1 MCCLPVFVILLITLTSAPSDALPKTRDDVPLASPHGGYNARRILQ--RRQGWCC 53
 ||||:||||| ||| | ||||| ||||:| | : ||: | |||
 RESULT 9
 AAY97119
 ID AAY97119 standard; Protein; 61 AA.
 XX
 AC AAY97119;
 XX
 DT 22-DEC-2000 (first entry)
 XX
 DE Tau conotoxin protein precursor.
 XX
 KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
 KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
 KW migraine; treatment; therapy.
 XX
 OS Conus geographus.
 XX
 PN WO200046371-A1.
 XX
 PD 10-AUG-2000.
 XX
 PF 04-FEB-2000; 2000WO-US03021.
 XX
 PR 04-FEB-1999; 99US-0118642.
 XX
 PA (UTAH) UNIV UTAH RES FOUND.
 PA (COGN-) COGNETIX INC.
 XX
 PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
 PI Jones RM;
 XX
 DR WPI: 2000-543489/49.
 DR N-PSDB; AAA51976.
 XX
 PT Novel analgesic short peptides from predatory cone snails for treating
 PT acute, chronic and neuropathic pain and migraines
 XX
 PS Claim 33; Page 18-19; 58pp; English.
 XX
 CC The venom of predatory cone snails comprises relatively small
 CC peptides which are targeted to various neuromuscular receptors
 CC and may be equivalent in their pharmacological diversity to the

```

Query Match      43.1%; Score 140; DB 21; Length 63;
Best Local Similarity 52.8%; Pred. No. 6e-09;
Matches 28; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 1 MRCLPVLIIILLITASAPGVVVLPTKEDDVPMSVYGNKSKILRGILRVCC 53
   |||:||||| ||| | ||:||||: |||: ||| : ||
Db 1 MRYLPVFVILLIIASIPSDTVOLKTKDDMPLASFGNGRRILRLMSNKRLLCC 53

RESULT 11
AAY97115
ID AAY97115 standard; Protein; 62 AA.
XX AC AAY97115;
XX DT 22-DEC-2000 (first entry)
XX DE Tau conotoxin protein precursor.
XX KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
XX KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
XX KW migraine; treatment; therapy.
XX OS Conus textile.
XX PN WO200046371-A1.
XX PD 10-AUG-2000.
XX PF 04-FEB-2000; 2000WO-US03021.
XX PR 04-FEB-1999; 99US-0118642.
XX PA (UTAH ) UNIV UTAH RES FOUND.
XX PA (COGN-) COGNETIX INC.
XX PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
XX PJ Jones RM;
XX DR WPI; 2000-543489/49.
XX DR N-PSDB; AAA51972.
XX PT Novel analgesic short peptides from predatory cone snails for treating
XX PT acute, chronic and neuropathic pain and migraines
XX PS Claim 33; Page 17; 58pp; English.
XX CC The venom of predatory cone snails comprises relatively small
XX CC peptides which are targetted to various neuromuscular receptors
XX CC and may be equivalent in their pharmacological diversity to the
XX CC alkaloids of plants or secondary metabolites of microorganisms.
XX CC Several peptides have been characterised from Conus venoms. These
XX CC include the alpha and mu conotoxins which target nicotinic
XX CC acetylcholine receptors and muscle sodium channels respectively.
XX CC Chronic or intractable pain as well as neuropathic pain is
XX CC currently treated with a range of analgesic compounds. It is
XX CC thought that the tau-conotoxins described could have applications
XX CC as analgesic drugs and could be used for treating such acute,
XX CC chronic and neuropathic pain as well as migraines.
XX SQ Sequence 62 AA;

Query Match      42.5%; Score 138; DB 21; Length 62;
Best Local Similarity 54.7%; Pred. No. 1e-08;
Matches 29; Conservative 8; Mismatches 14; Indels 2; Gaps 1;

QY 1 MRCLPVLIIILLITASAPGVVVLPTKEDDVPMSVYGNKSKILRGILRVCC 53
   |||:||||| ||| | ||:||||: |||: ||| : ||
Db 1 MCCLPVFVILLIIASAPSDVAPQPKTKDDVPLAPLHDNAKSAHQHL--NQRC 51

RESULT 12

```

```

AAY97125
ID AAY97125 standard; Protein; 64 AA.
XX AC AAY97125;
XX DT 22-DEC-2000 (first entry)
XX DE Tau conotoxin protein precursor.
XX KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
XX KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
XX KW migraine; treatment; therapy.
XX OS Conus caracteristicus.
XX PN WO200046371-A1.
XX PD 10-AUG-2000.
XX PF 04-FEB-2000; 2000WO-US03021.
XX PR 04-FEB-1999; 99US-0118642.
XX PA (UTAH ) UNIV UTAH RES FOUND.
XX PA (COGN-) COGNETIX INC.
XX PI Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
XX PJ Jones RM;
XX DR WPI; 2000-543489/49.
XX DR N-PSDB; AAA51982.
XX PT Novel analgesic short peptides from predatory cone snails for treating
XX PT acute, chronic and neuropathic pain and migraines
XX PS Claim 33; Page 21; 58pp; English.
XX CC The venom of predatory cone snails comprises relatively small
XX CC peptides which are targetted to various neuromuscular receptors
XX CC and may be equivalent in their pharmacological diversity to the
XX CC alkaloids of plants or secondary metabolites of microorganisms.
XX CC Several peptides have been characterised from Conus venoms. These
XX CC include the alpha and mu conotoxins which target nicotinic
XX CC acetylcholine receptors and muscle sodium channels respectively.
XX CC Chronic or intractable pain as well as neuropathic pain is
XX CC currently treated with a range of analgesic compounds. It is
XX CC thought that the tau-conotoxins described could have applications
XX CC as analgesic drugs and could be used for treating such acute,
XX CC chronic and neuropathic pain as well as migraines.
XX SQ Sequence 64 AA;

Query Match      42.0%; Score 136.5; DB 21; Length 64;
Best Local Similarity 53.3%; Pred. No. 1.6e-08;
Matches 32; Conservative 8; Mismatches 13; Indels 7; Gaps 2;

QY 1 MRCLPVLIIILLITASAPGVVVLPTKEDDVPMSVYGNKSKILRGILR---NGVCCGYKL 57
   |||:||||| ||| | ||:||||: |||: ||| : ||
Db 1 MRCLPVLIIILLITASAPGVDAQPKTKYNAPLTLSDHNAK---GILQHNKRCPPRL 56

RESULT 13
AAY97121
ID AAY97121 standard; Protein; 67 AA.
XX AC AAY97121;
XX DT 22-DEC-2000 (first entry)
XX DE Tau conotoxin protein precursor.
XX KW Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
XX KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;

```

KW migraine; treatment; therapy.

OS Conus textile.

PN WO200046371-A1.

XX PD 10-AUG-2000.

XX PF 04-FEB-2000; 2000WO-US03021.

XX PR 04-FEB-1999; 99US-0118642.

XX PA (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
PI Jones RM;

XX WPI; 2000-543489/49.

XX DR N-PSDB; AAA51978.

XX Novel analgesic short peptides from predatory cone snails for treating
acute, chronic and neuropathic pain and migraines

PS Claim 33; Page 19-20; 58pp; English.

XX The venom of predatory cone snails comprises relatively small
peptides which are targetted to various neuromuscular receptors
and may be equivalent in their pharmacological diversity to the
alkaloids of plants or secondary metabolites of microorganisms.
CC Several peptides have been characterised from Conus venoms. These
include the alpha and mu conotoxins which target nicotinic
acetylcholine receptors and muscle sodium channels respectively.
CC Chronic or intractable pain as well as neuropathic pain is
currently treated with a range of analgesic compounds. It is
thought that the tau-conotoxins described could have applications
as analgesic drugs and could be used for treating such acute,
chronic and neuropathic pain as well as migraines.

XX Sequence 67 AA;

Query Match 40.0%; Score 130; DB 21; Length 67;

Best Local Similarity 50.8%; Pred. No. 9.4e-08;

Matches 30; Conservative 5; Mismatches 18; Indels 6; Gaps 1;

QY 1 MRCPLVLIILLTASAPGVVLPKTEDDVPMSSVYGCKSI-----LRGILRNGVCC 53

||| || ||||| |||| ||:|||||: || ||: || ||

Db 1 MRCFPVFIILLIASAPCFDARTKTDVPLSLRDNLKRTIRLRNGCCDGGCC 59

RESULT 14

AAAY97127

ID AAY97127 standard; Protein; 62 AA.

XX AC AAY97127;

XX DT 22-DEC-2000 (first entry)

XX DE Tau conotoxin protein precursor.

XX Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
KW migraine; treatment; therapy.

XX OS Conus quercinus.

XX PN WO200046371-A1.

XX PD 10-AUG-2000.

XX PF 04-FEB-2000; 2000WO-US03021.

XX PR 04-FEB-1999; 99US-0118642.

XX (UTAH) UNIV UTAH RES FOUND.
PA (COGN-) COGNETIX INC.

XX Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
PI Jones RM;

XX WPI; 2000-543489/49.

XX DR N-PSDB; AAA51984.

XX Novel analgesic short peptides from predatory cone snails for treating
acute, chronic and neuropathic pain and migraines

PS Claim 33; Page 22; 58pp; English.

XX The venom of predatory cone snails comprises relatively small
peptides which are targetted to various neuromuscular receptors
and may be equivalent in their pharmacological diversity to the
alkaloids of plants or secondary metabolites of microorganisms.
CC Several peptides have been characterised from Conus venoms. These
include the alpha and mu conotoxins which target nicotinic
acetylcholine receptors and muscle sodium channels respectively.
CC Chronic or intractable pain as well as neuropathic pain is
currently treated with a range of analgesic compounds. It is
thought that the tau-conotoxins described could have applications
as analgesic drugs and could be used for treating such acute,
chronic and neuropathic pain as well as migraines.

XX Sequence 62 AA;

Query Match 39.4%; Score 128; DB 21; Length 62;

Best Local Similarity 46.6%; Pred. No. 1.5e-07;

Matches 27; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

QY 1 MRCPLVLIILLTASAPGVVLPKTEDDVPMSSVYGCKSIILRNGVCCGYKLC 58

||| || ||||| |||| ||:|||||: || ||: || ||

Db 1 MRCFPVFIILLSPSAPSDVAPHTKDDVPOASLHDDAKRTLQVPMKRGCCAMLTC 58

RESULT 15

AAAY97120

ID AAY97120 standard; Protein; 67 AA.

XX AC AAY97120;

XX DT 22-DEC-2000 (first entry)

XX DE Tau conotoxin protein precursor.

XX Conotoxin; chronic pain; neuropathic pain; acetylcholine; receptor;
KW antagonist; analgesic; drug; peptide; tau; cone snail; Conus; venom;
KW migraine; treatment; therapy.

XX OS Conus textile.

XX PN WO200046371-A1.

XX PD 10-AUG-2000.

XX PF 04-FEB-2000; 2000WO-US03021.

XX PR 04-FEB-1999; 99US-0118642.

XX PA (UTAH) UNIV UTAH RES FOUND.

XX (COGN-) COGNETIX INC.

XX Walker C, Shetty R, Olivera BM, Hooper D, Jacobsen R, Steele D;
PI Jones RM;

XX WPI; 2000-543489/49.

XX DR N-PSDB; AAA51977.

XX Novel analgesic short peptides from predatory cone snails for treating

PT acute, chronic and neuropathic pain and migraines
XX
PS Claim 33; Page 19; 58pp; English.
XX
CC The venom of predatory cone snails comprises relatively small
CC peptides which are targetted to various neuromuscular receptors
CC and may be equivalent in their pharmacological diversity to the
CC alkaloids of plants or secondary metabolites of microorganisms.
CC Several peptides have been characterised from Conus venoms. These
CC include the alpha and mu conotoxins which target nicotinic
CC acetylcholine receptors and muscle sodium channels respectively.
CC Chronic or intractable pain as well as neuropathic pain is
CC currently treated with a range of analgesic compounds. It is
CC thought that the tau-conotoxins described could have applications
CC as analgesic drugs and could be used for treating such acute,
CC chronic and neuropathic pain as well as migraines.
XX
SQ Sequence 67 AA;
Query Match 39.1%; Score 127; DB 21; Length 67;
Best Local Similarity 54.7%; Pred. No. 2.1e-07;
Matches 29; Conservative 4; Mismatches 20; Indels 0; Gaps 0;
QY 1 MRCLPVLIIILLITASAPGVVLPKTEDDVPMSVYGNCKSILRLNGVCC 53
||| || ||||| |||| ||:||||: ||: ||: ||
Db 1 MRCFPVFIILLIIASAPCFDARTKDDDDVPLSLRDNLKRTIRLRLNRECC 53

Search completed: January 29, 2003, 09:30:59
Job time : 62 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:29:32 ; Search time 3.10227 seconds
(without alignments)
84.558 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCGXLLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	61.6	24	8	US-08-775-765-8
2	44	60.3	538	10	US-09-779-307-2
3	44	60.3	538	10	US-09-779-307-11
4	44	60.3	547	10	US-09-779-307-12
5	44	60.3	547	10	US-09-779-307-13
6	42	57.5	1200	10	US-09-826-508-3
7	42	57.5	3907	9	US-10-029-217A-24
8	39	53.4	31	10	US-09-894-882-461
9	39	53.4	39	10	US-09-894-882-248
10	39	53.4	39	10	US-09-894-882-497
11	39	53.4	41	8	US-08-969-137-1
12	39	53.4	41	8	US-08-969-137-4
13	39	53.4	67	10	US-09-894-882-247
14	39	53.4	720	10	US-09-756-071B-20
15	39	53.4	1300	9	US-10-174-590-269
16	39	53.4	1300	9	US-10-176-758-269
17	39	53.4	1300	9	US-10-175-737-269
18	39	53.4	1300	12	US-10-052-586-269
19	39	53.4	1422	10	US-09-735-933-1

20	39	53.4	1435	9	US-10-174-590-581
21	39	53.4	1435	9	US-10-176-758-581
22	39	53.4	1435	9	US-10-175-737-581
23	39	53.4	1435	12	US-10-052-586-581
24	39	53.4	2211	9	US-10-096-961-1
25	38	52.1	36	10	US-09-894-882-339
26	38	52.1	36	10	US-09-894-882-443
27	38	52.1	39	10	US-09-894-882-342
28	38	52.1	39	10	US-09-894-882-448
29	38	52.1	72	10	US-09-894-882-338
30	38	52.1	72	10	US-09-894-882-341
31	38	52.1	187	10	US-09-811-284-178
32	38	52.1	212	9	US-09-738-626-5604
33	38	52.1	1497	10	US-09-060-854B-2
34	37	50.7	40	10	US-09-894-882-382
35	37	50.7	73	10	US-09-894-882-98
36	37	50.7	81	10	US-09-864-761-34804
37	37	50.7	81	10	US-09-864-761-41805
38	37	50.7	585	10	US-09-817-774-24
39	37	50.7	1404	10	US-09-811-045A-1
40	36.5	50.0	38	10	US-09-864-761-38941
41	36	49.3	26	8	US-08-775-765-9
42	36	49.3	32	9	US-09-749-637A-108
43	36	49.3	65	10	US-09-864-761-41688
44	36	49.3	210	9	US-10-114-893-107
45	36	49.3	441	12	US-10-115-406-4

ALIGNMENTS

RESULT 1
US-08-775-765-8
; Sequence 8, Application US/08775765C
; Patent No. US20010010821A1
; GENERAL INFORMATION:
; APPLICANT: Kelleher, Dermot
; APPLICANT: Windle, Henry
; APPLICANT: Byrne, William
; APPLICANT: McGarrus, Ross
; TITLE OF INVENTION: Helicobacter Proteins and Vaccines
; FILE REFERENCE: 08/775 765
; CURRENT APPLICATION NUMBER: US/08/775,765C
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IE 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IE 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-775-765-8

Query Match 61.6%; Score 45; DB 8; Length 24;
Best Local Similarity 50.0%; Pred. No. 2.6;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 GVC CGXLLCHXC 13
Db 6 GC CGCTTC TC 17

RESULT 2
US-09-779-307-2
; Sequence 2, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond

```
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-2

Query Match      60.3%; Score 44; DB 10; Length 538;
Best Local Similarity 50.0%; Pred. No. 38;
Matches      8; Conservative      0; Mismatches      4; Indels      4; Gaps      1;

Qy      2  GVCC-----GXXLCHXC 13
Db      508  GCCCFPLDGHLLCHGC 523

RESULT 3
US-09-779-307-11
; Sequence 11, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-11

Query Match      60.3%; Score 44; DB 10; Length 538;
Best Local Similarity 50.0%; Pred. No. 38;
Matches      8; Conservative      0; Mismatches      4; Indels      4; Gaps      1;

Qy      2  GVCC-----GXXLCHXC 13
Db      508  GCCCFPLDGHLLCHGC 523

RESULT 4
US-09-779-307-12
; Sequence 12, Application US/09779307
```

```
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-779-307-12

Query Match      60.3%; Score 44; DB 10; Length 547;
Best Local Similarity 50.0%; Pred. No. 39;
Matches      8; Conservative      0; Mismatches      4; Indels      4; Gaps      1;

Qy      2  GVCC---GXXLCHXC 13
Db      517  GCCCFPLDGHLLCHGC 532

RESULT 5
US-09-779-307-13
; Sequence 13, Application US/09779307
; Patent No. US20020137675A1
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-779-307-13

Query Match      60.3%; Score 44; DB 10; Length 547;
Best Local Similarity 50.0%; Pred. No. 39;
Matches      8; Conservative      0; Mismatches      4; Indels      4; Gaps      1;

Qy      2  GVCC---GXXLCHXC 13
Db      517  GCCCFPLDGHLLCHGC 532
```

```
RESULT 6
US-09-826-508-3
; Sequence 3, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; FILE REFERENCE: GP-70744USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1200
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-826-508-3

Query Match      57.5%; Score 42; DB 10; Length 1200;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      2  GVCCGXXLCHXC 13
Db      300  GACCGCTTCCAC 311

RESULT 7
US-10-029-217A-24
; Sequence 24, Application US/10029217A
; Patent No. US20020164735A1
; GENERAL INFORMATION:
; APPLICANT: WANG, DA-ZHI
; APPLICANT: OLSON, ERIC N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO A CARDIAC-SPECIFIC
; FILE REFERENCE: UTSD:695US
; CURRENT APPLICATION NUMBER: US/10/029,217A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/257,761
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-029-217A-24

Query Match      57.5%; Score 42; DB 9; Length 3907;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy      2  GVCCGXXLCHXC 13
Db      437  GCCCGAGTCAGC 448

RESULT 8
US-09-894-882-461
; Sequence 461, Application US/09894882
; Patent No. US20020102607A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Walker, Craig S.
; APPLICANT: Shetty, Reshma
; APPLICANT: Jimenez, Elsie C.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Watkins, Maren
```

```
; APPLICANT: Jones, Robert M.
; APPLICANT: Shen, Greg S.
; TITLE OF INVENTION: I-Superfamily Conotoxins
; FILE REFERENCE: 2314-238
; CURRENT APPLICATION NUMBER: US/09/894,882
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/243,410
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/246,581
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/247,714
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/264,256
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 461
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Conus emaciatus
US-09-894-882-461

Query Match      53.4%; Score 39; DB 10; Length 31;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

Qy      2  GVCCGX-XLCH 11
Db      17  GICCGTCRNVC 28

RESULT 9
US-09-894-882-248
; Sequence 248, Application US/09894882
; Patent No. US20020102607A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Walker, Craig S.
; APPLICANT: Shetty, Reshma
; APPLICANT: Jimenez, Elsie C.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: Shen, Greg S.
; TITLE OF INVENTION: I-Superfamily Conotoxins
; FILE REFERENCE: 2314-238
; CURRENT APPLICATION NUMBER: US/09/894,882
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/243,410
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/246,581
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/247,714
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/264,256
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 248
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Conus emaciatus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(39)
; OTHER INFORMATION: Xaa at residues 3, 4, 10 and 13 is Pro or hydroxy-Pro; Xaa at
; OTHER INFORMATION: idue 39 is Glu or gamma-carboxy-Glu; Xaa at residue 16 is Trp
```

```
; OTHER INFORMATION: bromo-Trp: Xaa at residues 7 and 11 is Tyr, 125I-Tyr, mono-iodo-Ty
; OTHER INFORMATION: yr, di-iodo-Tyr, O-sulpho-Tyr or O-phospho-Ty
US-09-894-882-488

Query Match      53.4%; Score 39; DB 10; Length 39;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY  2  GVCCGX--XLCH 11
Db  17  GICCGTCRNCH 28

RESULT 10
US-09-894-882-497
; Sequence 497, Application US/09894882
; Patent No. US20020102607A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Walker, Craig S.
; APPLICANT: Shetty, Reshma S.
; APPLICANT: Jimenez, Elsie C.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: Shen, Greg S.
; TITLE OF INVENTION: I-Superfamily Conotoxins
; FILE REFERENCE: 2314-238
; CURRENT APPLICATION NUMBER: US/09/894,882
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/243,410
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/246,581
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/247,714
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/264,256
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 497
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Conus emaciatus
US-09-894-882-497

Query Match      53.4%; Score 39; DB 10; Length 39;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY  2  GVCCGX--XLCH 11
Db  17  GICCGTCRNCH 28

RESULT 11
US-08-969-137-1
; Sequence 1, Application US/08969137
; Patent No. US20010018207A1
; GENERAL INFORMATION:
; APPLICANT: KANDEL, ERIC
; APPLICANT: MAYFORD, MARK
; TITLE OF INVENTION: DNA REGULATORY ELEMENT FOR THE
; TELECOMMUNICATION INFORMATION: 0575/52776
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER AND DUNHAM
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK

; OTHER INFORMATION: bromo-Trp: Xaa at residues 7 and 11 is Tyr, 125I-Tyr, mono-iodo-Ty
; OTHER INFORMATION: yr, di-iodo-Tyr, O-sulpho-Tyr or O-phospho-Ty
US-09-894-882-488

Query Match      53.4%; Score 39; DB 10; Length 39;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY  2  GVCCGX--XLCH 11
Db  17  GICCGTCRNCH 28

RESULT 12
US-08-969-137-4
; Sequence 4, Application US/08969137
; Patent No. US20010018207A1
; GENERAL INFORMATION:
; APPLICANT: KANDEL, ERIC
; APPLICANT: MAYFORD, MARK
; TITLE OF INVENTION: DNA REGULATORY ELEMENT FOR THE
; TELECOMMUNICATION INFORMATION: 0575/52776
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER AND DUNHAM
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/969,137
; FILING DATE: 12-NOV-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)2276821
; TELEFAX: (212)3910525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-969-137-4

Query Match 53.4%; Score 39; DB 8; Length 41;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGX LCHXC 13
Db 15 GCCCGTTCGACC 26

RESULT 13
US-09-894-882-247
; Sequence 247, Application US/09894882
; Patent No. US20020102607A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Walker, Craig S.
; APPLICANT: Shetty, Reshma
; APPLICANT: Jimenez, Elsie C.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Watkins, Maren
; APPLICANT: Jones, Robert M.
; APPLICANT: Shen, Greg S.
; TITLE OF INVENTION: I-Superfamily Conotoxins
; FILE REFERENCE: 2314-238
; CURRENT APPLICATION NUMBER: US/09/894,882
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/243,410
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/246,581
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/247,714
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/264,256
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 247
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Conus emaciatus
US-09-894-882-247

Query Match 53.4%; Score 39; DB 10; Length 67;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 2 GVC CGX --XLCH 11
Db 45 GICCGTCRNCH 56

RESULT 14
US-09-756-071B-20
; Sequence 20, Application US/09756071B
; Patent No. US20020052307A1
; GENERAL INFORMATION:
; APPLICANT: Tryggvason, Karl
; APPLICANT: kallunki, Pekka
; APPLICANT: Pyke, Charles
; TITLE OF INVENTION: Laminin Chains: Diagnostic Uses
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fay Sharpe Fagan Minnich & McKee
```

```
; STREET: 1100 Superior Ave, Suite 700
; CITY: Cleveland
; STATE: Ohio
; COUNTRY: USA
; ZIP: 44114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,071B
; FILING DATE: 08-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/663,147
; FILING DATE: 150-September 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Minnich, Richard, J.
; REGISTRATION NUMBER: 24,175
; REFERENCE/DOCKET NUMBER: TRV 20014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 216-861-5582
; TELEFAX: 216-241-1666
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-756-071B-20

Query Match 53.4%; Score 39; DB 10; Length 720;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGX LCHXC 13
Db 690 GCCCGGCACCGC 701

RESULT 15
US-10-174-590-269
; Sequence 269, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 269
; LENGTH: 1300
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-269

Query Match 53.4%; Score 39; DB 9; Length 1300;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
```

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13

Db 794 GGCCGAGCCCGC 805

Search completed: January 29, 2003, 09:34:07
Job time : 4.10227 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 10.1932 Seconds
(without alignments)
262.785 Million cell updates/sec

Title: US-09-580-201a-2
Perfect score: 73
Sequence: 1 NGVCCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	64.4	326	Q8RA96	Q8RA96 thermoanaer
2	46	63.0	258	O36453	O36453 lymantria d
3	46	63.0	514	Q95Q73	Q95Q73 caenorhabdi
4	45	61.6	244	Q9IBQ9	Q9IBQ9 spodoptera
5	44	60.3	121	Q9CZ7	Q9CZ7 mus musculus
6	44	60.3	372	Q9ND4	Q9ND4 rattus norv
7	44	60.3	538	Q96IF1	Q96IF1 homo sapien
8	44	60.3	547	P97472	P97472 mus musculus
9	44	60.3	547	Q91XC0	Q91XC0 mus musculus
10	43	58.9	104	Q9REV9	Q9REV9 pseudomonas
11	43	58.9	310	Q9WUG1	Q9WUG1 cavia porce
12	43	58.9	563	Q9N3Y4	Q9N3Y4 caenorhabdi
13	43	58.9	746	Q9TUT3	Q9TUT3 oryctolagus
14	43	58.9	746	Q9GKE7	Q9GKE7 sus scrofa
15	43	58.9	746	Q9P66	Q9P66 cavia porce
16	43	58.9	808	Q9PWK7	Q9PWK7 xenopus lae

17	43	58.9	808	13	O13080	O13080 xenopus lae
18	42	57.5	105	2	O85226	O85226 pseudomonas
19	42	57.5	151	12	O10614	O10614 helicoverpa
20	42	57.5	171	16	O8XFK4	O8XFK4 salmonella
21	42	57.5	171	16	O8XE86	O8XE86 escherichia
22	42	57.5	234	12	O8QL66	O8QL66 mamestra co
23	42	57.5	244	5	O62463	O62463 caenorhabdi
24	42	57.5	244	5	O62464	O62464 caenorhabdi
25	42	57.5	285	12	O9H38	O9H38 helicoverp
26	42	57.5	285	12	O8V5X4	O8V5X4 helicoverpa
27	42	57.5	289	12	Q91BK7	Q91BK7 spodoptera
28	42	57.5	322	10	Q39896	Q39896 glycine max
29	42	57.5	326	10	Q39895	Q39895 glycine max
30	42	57.5	332	5	O18012	O18012 caenorhabdi
31	42	57.5	446	4	Q96B03	Q96B03 homo sapien
32	42	57.5	954	4	Q9ULG5	Q9ULG5 homo sapien
33	42	57.5	1166	4	Q9P2R3	Q9P2R3 homo sapien
34	42	57.5	1184	11	O54807	O54807 mus musculu
35	42	57.5	1441	10	Q9LK63	Q9LK63 arabidopsis
36	42	57.5	1466	10	O8VZ24	O8VZ24 arabidopsis
37	41	56.2	179	16	Q9HVH1	Q9HVH1 pseudomonas
38	41	56.2	263	16	Q92NF2	Q92NF2 rhizobium m
39	41	56.2	272	5	Q9ULS9	Q9ULS9 caenorhabdi
40	41	56.2	328	16	Q97H26	Q97H26 clostridium
41	41	56.2	338	5	Q9USC2	Q9USC2 caenorhabdi
42	41	56.2	350	5	Q9N8U4	Q9N8U4 trypanosoma
43	41	56.2	362	5	Q9U483	Q9U483 neospora ca
44	41	56.2	363	5	Q9N5Y3	Q9N5Y3 caenorhabdi
45	41	56.2	465	4	O60260	O60260 homo sapien

ALIGNMENTS

RESULT 1

- Q8RA96 PRELIMINARY; PRT; 326 AA.
- AC Q8RA96;
- DT 01-JUN-2002 (TREMBLrel. 21, Created)
- DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
- DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
- DE 2-polyphenylphenol hydroxylase and related flavodoxin oxidoreductases.
- GN UBIB3 OR TTEL329.
- OS Thermoanaerobacter tengcongensis.
- OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
- OC Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
- OX NCBI_TaxID=119072;
- RN [1]
- RP SEQUENCE FROM N.A.
- RC STRAIN=MB4T / JCM11007;
- RX MEDLINE=21992816; PubMed=11997336;
- RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J., Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L., Tan H., Chen R., Wang J., Yu J., Yang H.;
- RT "A complete sequence of T. tengcongensis genome.";
- RL Genome Res. 12:689-700(2002).
- DR EMBL; AE013093; AAM24553.1; -.
- KW Complete proteome.
- SQ SEQUENCE 326 AA; 36899 MW; 117C97B98BA8A63B CRC64;

Query Match 64.4%; Score 47; DB 16; Length 326;
Best Local Similarity 46.2%; Pred. No. 2.4; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 5;

Qy 1 NGVCCGXXLCHXC 13
| : | | | : |
Db 285 NKICCGEGICGSC 297

RESULT 2
O36453 PRELIMINARY; PRT; 258 AA.
ID O36453

AC O36453; Q9YMW3;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Immediate early 0 protein (Immediate early transactivator 0).
 GN IE-0.
 OS Lymantria dispar multicausid nuclear polyhedrosis virus (LdMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97445058; PubMed=9300047;
 RA Pearson M.N., Rohrmann G.F.;
 RA "Splicing is required for transactivation by the immediate early gene
 1 of the Lymantria dispar multinucleocapsid nuclear polyhedrosis
 virus.";
 RT Virology 253:17-34(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124785; PubMed=9887315;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J., Rohrmann G.F.;
 RA Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF006656; AAC38234.1; -;
 DR EMBL; AF081810; AAC70206.1; -;
 DR InterPro; IPR001841; Znf_ring.
 DR SMART; SM00184; RING; 1.
 SQ SEQUENCE 258 AA; 29395 MW; AC564CDF92282BAD CRC64;
 Query Match 63.0%; Score 46; DB 12; Length 258;
 Best Local Similarity 54.5%; Pred. No. 2.9;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 3 VCCGXXLCHXC 13
 DB 208 VCCGYRVCNAC 218
 ||||| :| :|
 RESULT 3
 Q95QT3
 ID Q95QT3 PRELIMINARY; PRT; 514 AA.
 AC Q95QT3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 58.6 kDa protein.
 GN C2861.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodirinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;
 RA Favello A.;
 RT "The sequence of C. elegans cosmid C28G1.";
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
 DR EMBL; U41026; AAL02448.1; -;
 DR InterPro; IPR000315; Znf_Bbox.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00643; zf-B_box; 1.
 DR Pfam; PF00097; zf-C3HC4; 2.
 DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_2.
 KW Hypothetical protein; Zinc-finger.
 SQ SEQUENCE 514 AA; 58600 MW; BC33388F0F599446 CRC64;
 Query Match 63.0%; Score 46; DB 5; Length 514;
 Best Local Similarity 63.6%; Pred. No. 4.9;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 3 VCCGXXLCHXC 13
 DB 263 VTCGHALCHKC 273
 |||||
 RESULT 4
 Q9IBQ9
 ID Q9IBQ9 PRELIMINARY; PRT; 244 AA.
 AC Q9IBQ9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ORF138 ie0.
 OS Spodoptera exigua nucleopolyhedrovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10454;
 RN [1]
 RP PARTIAL SEQUENCE FROM N.A.
 RX MEDLINE=97437494; PubMed=9292027;
 RA van Strien E.A., Faktor O., Hu Z.H., Zuidema D., Goldbach R.W.,
 RA Viak J.M.;
 RT "Baculoviruses contain a gene for the large subunit of ribonucleotide
 reductase.";
 RL J. Gen. Virol. 78:2365-2377(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20036646; PubMed=10567663;
 RA IJkel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
 RA Goldbach R.W., Viak J.M.;
 RT "Sequence and organization of the spodoptera exigua multicausid
 nucleopolyhedrovirus genome.";
 RL J. Gen. Virol. 80:3289-3304(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Zuidema D.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA IJkel W.F.J., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
 RA Goldbach R.W., Viak J.M.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF169823; AAF33667.1; -;
 DR InterPro; IPR001841; Znf_ring.
 DR SMART; SM00184; RING; 1.
 SQ SEQUENCE 244 AA; 28719 MW; 1F7662E837A866DB CRC64;
 Query Match 61.6%; Score 45; DB 12; Length 244;
 Best Local Similarity 60.0%; Pred. No. 4;


```

Matches      6;  Conservative      1;  Mismatches      3;  Indels      0;  Gaps      0;

QY      4  CCGXXLCHXC 13
      ||| ||| |
Db      199  CCGYNLCYVC 208

RESULT 5
Q9CZW7
ID      Q9CZW7      PRELIMINARY;      PRT;      121 AA.
AC      Q9CZW7;
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE      10 days embryo cDNA, RIKEN full-length enriched library,
DE      clone:2610509F17, full insert sequence.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX      MEDLINE=21085660; PubMed=11217851;
RA      Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA      Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA      Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA      Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA      Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA      Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA      Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA      Schriml L.M., StaUBL F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA      Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA      Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA      Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA      Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA      Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA      Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA      Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA      Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA      Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA      Hayashizaki Y.;
RT      "Functional annotation of a full-length mouse cDNA collection.";
RL      Nature 409:685-690(2001).
CC      -!- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC      IONS.
DR      EMBL; AK012077; BAB28013.1; -.
DR      HSSP; P32965; 1CTL.
DR      InterPro; IPR001781; LIM.
DR      Pfam; PF00412; LIM; 2.
DR      ProDom; PD000094; LIM; 2.
DR      SMART; SM00132; LIM; 2.
DR      PROSITE; PS50023; LIM_DOMAIN_2; 2.
KW      LIM domain; Metal-binding; Zinc.
SQ      SEQUENCE 121 AA; 13774 MW; 5DB20785DE5F74DE CRC64;

Query Match      60.3%; Score 44; DB 11; Length 121;
Best Local Similarity 50.0%; Pred. No. 3.4;
Matches      8;  Conservative      0;  Mismatches      4;  Indels      4;  Gaps      1;

QY      2  GVCC----GXLLCHXC 13
      | | | | | | | |
Db      91  GCCCFLDGHLLCHGC 106

RESULT 6
Q99ND4
ID      Q99ND4      PRELIMINARY;      PRT;      372 AA.
AC      Q99ND4;
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT      01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE      Ajuba protein (Fragment).
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Thesis (2000), University College London, University of London,
RC      London, United Kingdom.
RL      -!- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC      IONS.
DR      EMBL; AJ306292; CAC28536.1; -.
DR      HSSP; Q05158; 1QLI.
DR      InterPro; IPR001781; LIM.
DR      Pfam; PF00412; LIM; 3.
DR      ProDom; PD000094; LIM; 3.
DR      SMART; SM00132; LIM; 3.
DR      PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR      PROSITE; PS50023; LIM_DOMAIN_2; 3.
KW      LIM domain; Metal-binding; Zinc.
FT      NON_TER      1
SQ      SEQUENCE 372 AA; 40210 MW; A6F109D84F41AFC1 CRC64;

Query Match      60.3%; Score 44; DB 11; Length 372;
Best Local Similarity 50.0%; Pred. No. 8.1;
Matches      8;  Conservative      0;  Mismatches      4;  Indels      4;  Gaps      1;

QY      2  GVCC----GXLLCHXC 13
      | | | | | | | |
Db      342  GCCCFLDGHLLCHGC 357

RESULT 7
Q96IF1
ID      Q96IF1      PRELIMINARY;      PRT;      538 AA.
AC      Q96IF1;
DT      01-DEC-2001 (TReMBLrel. 19, Created)
DT      01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE      Similar to ajuba.
DE      Similar to ajuba.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RA      Strausberg R.;
RL      Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC      IONS.
DR      EMBL; BC007580; AAH07580.1; -.
DR      InterPro; IPR001781; LIM.
DR      InterPro; IPR003662; sub_transporter.
DR      Pfam; PF00412; LIM; 3.
DR      ProDom; PD000094; LIM; 3.
DR      PROSITE; PS00478; LIM_DOMAIN_1; UNKNOWN_2.
DR      PROSITE; PS50023; LIM_DOMAIN_2; 3.
DR      PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW      LIM domain; Metal-binding; Zinc.
SQ      SEQUENCE 538 AA; 56933 MW; 1E0DFA0336976A3C CRC64;

Query Match      60.3%; Score 44; DB 4; Length 538;
Best Local Similarity 50.0%; Pred. No. 11;
Matches      8;  Conservative      0;  Mismatches      4;  Indels      4;  Gaps      1;

QY      2  GVCC----GXLLCHXC 13

```

```
Db 508 GCCCFPLDGHLLCHGC 523
      ||| | | ||| |
      1 11 1 111 1
Query Match 60.3%; Score 44; DB 11; Length 547;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
DR PRODom; PD000094; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; UNKNOWN_2.
DR PROSITE; PS0023; LIM_DOMAIN_2; 3.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 547 AA; 57919 MW; C326772AC1C155B2 CRC64;

RESULT 8
P97472
ID P97472 PRELIMINARY; PRT; 547 AA.
AC P97472;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Ajuba.
GN JUB OR AJUBA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99263009; PubMed=10330178;
RA Goyal R.K., Lin P., Kanungo J., Payne A.S., Muslin A.J.,
RA Longmore G.D.;
RT "Ajuba, a novel LIM protein, interacts with Grb2, augments mitogen-
RT activated protein kinase activity in fibroblasts, and promotes meiotic
RT maturation of Xenopus oocytes in a Grb2- and Ras-dependent manner.";
RL Mol. Cell. Biol. 19:4379-4389(1999).
CC -!- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL; J09776; AAB38287.1; -.
DR HSSP; Q05158; IQLI.
DR MGD; MGI:1341886; Jub.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.
DR ProDom; PD000094; LIM; 3.
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS0023; LIM_DOMAIN_2; 3.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 547 AA; 57901 MW; C326772AC1C441B2 CRC64;

Query Match 60.3%; Score 44; DB 11; Length 547;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;
Qy 2 GVCC---GXXLCHXC 13
      ||| | | ||| |
      1 11 1 111 1
Db 517 GCCCFPLDGHLLCHGC 532

RESULT 9
Q91XC0
ID Q91XC0 PRELIMINARY; PRT; 547 AA.
AC Q91XC0;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Similar to ajuba.
GN JUB
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Gland;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS.
DR EMBL; BC011035; AAH11035.1; -.
DR MGD; MGI:1341886; Jub.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 3.

Query Match 58.9%; Score 43; DB 16; Length 104;
Best Local Similarity 46.2%; Pred. No. 4.4;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 NGVCCGXXLCHXC 13
      :| | | | |
      56 SGAFCGMGVCHCC 68

Db 56 SGAFCGMGVCHCC 68

RESULT 11
Q9WUG1
ID Q9WUG1 PRELIMINARY; PRT; 310 AA.
AC Q9WUG1;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Chloride channel Clc-5 (Fragment).
OS Cavia porcellus (Guinea pig).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE EPITHELIUM;
RA Cid L.P., Salinas O., Sepulveda F.V.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF133215; AAD33600.1; -.
DR InterPro; IPR001807; Cl-channel_volt.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
FT NON_TER 1
FT NON_TER 310
SQ SEQUENCE 310 AA; 34691 MW; 15419E709DB531EF CRC64;

Query Match 58.9%; Score 43; DB 11; Length 310;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
Db 165 CCGNILCH 172

RESULT 12
Q9N3Y4 PRELIMINARY; PRT; 563 AA.
AC Q9N3Y4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical 63.7 kDa protein.
GN Y42H9AR.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nemata; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Tin-Wollam A., Wohldmann P.;
RT "The sequence of C. elegans cosmid Y42H9AR.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024777; AAF60562.3; -.
DR InterPro; IPR000822; Znf_C2H2.
DR InterPro; IPR000306; Znf_FIVE.
DR Pfam; PF01363; FIVE; 1.
DR SMART; SM00064; FIVE; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
KW DNA-binding; Hypothetical protein; Metal-binding; Zinc-finger.
SQ SEQUENCE 563 AA; 63682 MW; 5864FEBD91432DD3 CRC64;

Query Match 58.9%; Score 43; DB 5; Length 563;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 CGXXLCHXC 13
```

```
Db 204 CGRVLCHSC 212
|| ||| |
RESULT 13
Q9TTU3 PRELIMINARY; PRT; 746 AA.
AC Q9TTU3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Chloride channel CLC-5.
GN CLCN5.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE; TISSUE=CORNEA EPITHELIA;
RA Rae J.L.;
RT "Ion Channels in Cornea Epithelia.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF195523; AAF06018.1; -.
DR InterPro; IPR000644; CBS_domain.
DR InterPro; IPR001807; Cl-channel_volt.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
SQ SEQUENCE 746 AA; 83206 MW; 667F2701C0BF006A CRC64;

Query Match 58.9%; Score 43; DB 6; Length 746;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
Db 220 CCGNILCH 227

RESULT 14
Q9GKE7 PRELIMINARY; PRT; 746 AA.
AC Q9GKE7;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Outwardly rectifying chloride channel.
GN CLC-5.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20545523; PubMed=10978325;
RA Dowland L.K., Luyckx V.A., Enck A.H., Leclercq B., Yu A.S.L.;
RT "Molecular Cloning and Characterization of an Intracellular Chloride
Channel in the Proximal Tubule Cell Line, LLC-PK1.";
RL J. Biol. Chem. 275:37765-37773(2000).
DR EMBL; AF274055; AAG29104.1; -.
DR InterPro; IPR000644; CBS_domain.
DR InterPro; IPR001807; Cl-channel_volt.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
SQ SEQUENCE 746 AA; 83123 MW; 655E3ED45FC61229 CRC64;

Query Match 58.9%; Score 43; DB 6; Length 746;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 4 CCGXXLCH 11
 ||| |||
 Db 220 CCGNILCH 227

RESULT 15

Q99P66 PRELIMINARY; PRT; 746 AA.
 AC Q99P66;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Chloride channel CLCN5.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cornejo I., Cid L.P., Sepulveda F.V.;
 RT "Cloning and intestinal expression of guinea pig CLC-5 chloride
 channel."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF326968; AAG49590.1; .
 DR InterPro: IPR000644; CBS_domain.
 DR InterPro: IPR001807; Cl-channel_volt.
 DR Pfam: PF00571; CBS; 2.
 DR Pfam: PF00654; voltage_CLC; 1.
 DR PRINTS: PR00762; CLCHANNEL.
 DR SMART: SM00116; CBS; 2.
 SQ SEQUENCE 746 AA; 83093 MW; 92FC8AFDC7D8D4A5 CRC64;

Query Match 58.9%; Score 43; DB 11; Length 746;
 Best Local Similarity 75.0%; Pred. NO. 20;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
 ||| |||
 Db 220 CCGNILCH 227

Search completed: January 29, 2003, 09:32:16
 Job time : 12.1932 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 2.95455 seconds
(without alignments)
182.496 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	66	90.4	61	1 CXL2_CONMR	P58808 conus marmo
2	60	82.2	13	1 CXL4_CONMR	P58810 conus marmo
3	54	74.0	11	1 CXL1_CONMR	P58807 conus marmo
4	50	68.5	12	1 CXL3_CONMR	P58809 conus marmo
5	45	61.6	400	1 V678_HSV11	Q00167 ictalurid h
6	44	60.3	1235	1 DNB1_HCMVA	P17147 human cytom
7	43	58.9	746	1 CLC5_HUMAN	P51795 homo sapien
8	43	58.9	746	1 CLC5_MOUSE	Q9wvd4 mus musculu
9	43	58.9	746	1 CLC5_RAT	P51796 rattus norv
10	42	57.5	172	1 PGPA_ECOLI	P18200 escherichia
11	41.5	56.8	768	1 ITB8_RABIT	P26013 oryctolagus
12	41	56.2	161	1 NEUV_CHICK	P24787 gallus gall
13	41	56.2	554	1 ARD1_RAT	P36407 rattus norv
14	41	56.2	574	1 ARD1_HUMAN	P36406 homo sapien
15	40	54.8	837	1 UBPA_DICDI	P54201 dictyosteli
16	39	53.4	55	1 M84C_DROME	Q01644 drosophila
17	39	53.4	144	1 NEU2_CAVPO	P10769 cavia porce
18	39	53.4	245	1 IE0_NPVOP	O10369 orgyia pseu
19	39	53.4	266	1 GNP1_GIALA	O97439 giardia lam
20	39	53.4	1160	1 DNB1_GMVC	P13215 simian cyto
21	38.5	52.7	3175	1 RPOA_SAV	P19811 equine arte
22	38	52.1	155	1 NEU4_CATCO	P16229 catostomus
23	38	52.1	465	1 RN15_HUMAN	O00635 homo sapien
24	38	52.1	1191	1 DNB1_MCMVS	P30672 murine cyto
25	37	50.7	96	1 FSPM_LYCES	P14903 lycopersico
26	37	50.7	125	1 NEU1_BOVIN	P01175 bos taurus
27	37	50.7	125	1 NEU1_SHEEP	P13389 ovis aries
28	37	50.7	196	1 V17_BPT7	P03781 bacterioph
29	37	50.7	259	1 PYRK_FUSNN	P58884 fusobacteri
30	37	50.7	392	1 FZD9_CHICK	O9ia02 gallus gall
31	36.5	50.0	1808	1 TENA_CHICK	P10039 gallus gall
32	36	49.3	49	1 S325_SEGFL	P58605 segestria f
33	36	49.3	88	1 V184_FOWPV	Q9J549 fowlpox vir

RESULT 1				
ID	CXL2_CONMR	STANDARD:	PRT:	61 AA.
AC	P58808;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lambda-conotoxin CMrVIB precursor (Chi-conotoxin MrIA) (Chi-MrIA)			
DE	(mr10a).			
OS	Conus marmoreus (Marble cone).			
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;			
OC	Neogastropoda; Conoidea; Conidae; Conus.			
OX	NCBI_TaxID=42752;			
RN	[1]			
RP	SEQUENCE FROM N.A., SEQUENCE OF 49-61, SYNTHESIS, AND MASS			
RP	SPECTROMETRY.			
RC	TISSUE=Venom duct, and Venom;			
RX	MEDLINE=20490660; PubMed=10900201;			
RA	McIntosh J.M., Corpus G.O., Layer R.T., Garrett J.E., Wagstaff J.D.,			
RA	Bulaj G., Vyazovkina A., Yoshikami D., Cruz L.J., Olivera B.M.;			
RT	"Isolation and characterization of a novel conus peptide with apparent			
RT	anticoagulative activity."			
RL	J. Biol. Chem. 275:32391-32397(2000).			
RN	[2]			
RP	SEQUENCE OF 49-61, AND MASS SPECTROMETRY.			
RC	TISSUE=Venom;			
RX	MEDLINE=20564325; PubMed=10988292;			
RA	Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,			
RA	Seow K.T., Bay B.-H.;			
RT	"Lambda-conotoxins, a new family of conotoxins with unique disulfide			
RT	pattern and protein folding. Isolation and characterization from the			
RT	venom of Conus marmoreus."			
RL	J. Biol. Chem. 275:39516-39522(2000).			
RN	[3]			
RP	SEQUENCE OF 49-61, SYNTHESIS, MASS SPECTROMETRY, AND STRUCTURE BY NMR.			
RC	TISSUE=Venom;			
RX	MEDLINE=21419681; PubMed=11528421;			
RA	Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,			
RA	Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,			
RA	Lewis R.J.;			
RT	"Two new classes of conopeptides inhibit the alpha1-adrenoceptor and			
RT	noradrenaline transporter."			
RL	Nat. Neurosci. 4:902-907(2001).			
CC	-!- FUNCTION: Inhibits the neuronal noradrenaline transporter.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- TISSUE SPECIFICITY: Expressed by the venom duct.			
CC	-!- PTM: Exists in two forms, due to cis-trans isomerization at His-			
CC	59-Hyp-60.			
CC	-!- MASS SPECTROMETRY: MW=1408.5; METHOD=Electrospray.			
CC	-!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.			
KW	Neurotoxin; Toxin; Hydroxylation; Signal.			
FT	SIGNAL	1	19	POTENTIAL.
FT	PROPEP	20	48	
FT	PEPTIDE	49	61	LAMBDA-CONOTOXIN CMrVIB.
FT	DISULFID	52	61	
FT	DISULFID	53	58	

34	36	49.3	131	1	NEU2_ANSAN	P19630 anser anser
35	36	49.3	186	1	DHML_METFL	Q50425 methylobaci
36	36	49.3	187	1	DHML_METME	Q59543 methylobaci
37	36	49.3	206	1	CLD5_RAT	O91kd6 rattus norv
38	36	49.3	218	1	CLD5_MOUSE	O51942 mus musculu
39	36	49.3	233	1	TPL_TREPA	P45685 treponema p
40	36	49.3	244	1	CSMI_CHLFE	O69888 chlorobium
41	36	49.3	250	1	YS85_CAEEL	Q09623 caenorhabdi
42	36	49.3	261	1	IE0_NPVAC	P41710 autographa
43	36	49.3	309	1	MAT1_HUMAN	P51948 homo sapien
44	36	49.3	309	1	MAT1_MOUSE	P51949 mus musculu
45	36	49.3	309	1	MAT1_XENLA	P51951 xenopus lae

ALIGNMENTS

```

FT MOD_RES 60 60 HYDROXYLATION.
SQ SEQUENCE 61 AA; 6499 MW; F4DE5B5A97EB8DBA CRC64;

Query Match 90.4%; Score 66; DB 1; Length 61;
Best Local Similarity 76.9%; Pred. No. 0.0025;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXLCXHC 13
   ||||| ||| |
Db 49 NGVCCGYKLCHPC 61

RESULT 2
CXLA_CONMR STANDARD; PRT; 13 AA.
AC P58810;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda/chi-conotoxin MrJB (Chi-MrJB).
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
RC TISSUE=Venom;
RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter."
RL Nat. Neurosci. 4:902-907(2001).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- PTM: Exists in two forms, due to cis-trans isomerization at His-
CC 11-Hyp-12.
CC -!- MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB; 1IEO; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT DISULFID 5 10
FT MOD_RES 12 12 HYDROXYLATION.
FT SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 82.2%; Score 60; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.0059;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVC CGXXLCXHC 13
   ||||| ||| |
Db 2 GVC CGYKLCHPC 13

RESULT 3
CXLI_CONMR STANDARD; PRT; 11 AA.
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMrVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;

```

```

RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus."
RL J. Biol. Chem. 275:39516-39522(2000).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT DISULFID 3 8
FT MOD_RES 10 10 HYDROXYLATION.
FT SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;

Query Match 74.0%; Score 54; DB 1; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 VCCGXLCXHC 13
   |||| ||| |
Db 1 VCCGYKLCHPC 11

RESULT 4
CXLI_CONMR STANDARD; PRT; 12 AA.
AC P58809;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMrX.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus."
RL J. Biol. Chem. 275:39516-39522(2000).
CC -!- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1262.77; MW_ERR=0.07; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 3 12
FT DISULFID 4 9
FT MOD_RES 11 11 HYDROXYLATION.
FT SEQUENCE 12 AA; 1251 MW; 277AAE242D2D5A2C8 CRC64;

Query Match 68.5%; Score 50; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 0.14;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GVC CGXXLCXHC 13
   |||| | | |
Db 1 GTCGVSFCYPC 12

RESULT 5
VG78_HSV11
ID VG78_HSV11 STANDARD; PRT; 400 AA.
AC Q00167;
DT 01-DEC-1992 (Rel. 24, Created)

```

DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical gene 78 zinc-binding protein.
GN 78.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Auburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davidson A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M75136; AAA88180.1; -.
DR PIR; D36794; ZBBE14.
KW Hypothetical protein; Zinc; Zinc-finger.
SQ SEQUENCE 400 AA; 44167 MW; D941DB7738B95CA9 CRC64;
Query Match 61.6%; Score 45; DB 1; Length 400;
Best Local Similarity 60.0%; Pred. No. 7.6;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 CCGXXLCHXC 13
||| |||
DB 133 CCGATLCDCS 142
RESULT 6
ID DNBI_HCMVA STANDARD; PRT; 1235 AA.
AC P17147;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major DNA-binding protein (MDBP).
GN UL57 OR DBP.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kouratides T., Martignetti J.A.,
RA Predie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC -----
DR EMBL; X17403; CAA35372.1; -.
DR PIR; S09820; QOBEM4.
DR InterPro: IPR000635; Viral_DNA_bind.
DR Pfam: PF00747; Viral_DNA_bp_1.
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein;
KW Early protein.
FT ZN_FING 467
FT ZN_FING 481 C2HC-TYPE.
SQ SEQUENCE 1235 AA; 133878 MW; 94E8D4F8D3BB2CB6 CRC64;
Query Match 60.3%; Score 44; DB 1; Length 1235;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 2 GVC--CGXXLCHXC 13
||| ||| |||
DB 470 GLCEACGGTCCHTC 483
RESULT 7
ID CLC5_HUMAN STANDARD; PRT; 746 AA.
AC P51795;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLCN5 OR CLCK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney.
RX MEDLINE=96121370; PubMed=8575751;
RA Fisher S.E., van Bakel I., Lloyd S.E., Pearce S.H.S.,
RA Thakker R.V., Craig I.W.;
RT "Cloning and characterization of CLCN5, the human kidney chloride
RT channel gene implicated in Dent disease (an X-linked hereditary
RT nephrolithiasis).";
RL Nephrolithiasis. 29:598-606(1995).
RN [2]
RP SEQUENCE OF 487-746 FROM N.A.
RC TISSUE=Kidney.
RX MEDLINE=95179126; PubMed=7874126;
RA Fisher S., Black G.C.M., Lloyd S.E., Hatchwell E., Wrong O.,
RA Thakker R.V., Craig I.W.;
RT "Isolation and partial characterization of a chloride channel gene
RT which is expressed in kidney and is a candidate for Dent's disease
RT (an X-linked hereditary nephrolithiasis).";
RL Hum. Mol. Genet. 3:2053-2059(1994).
RN [3]
RP TISSUE SPECIFICITY.
RC TISSUE=Vascular smooth muscle, and Aortic endothelium;
RX MEDLINE=99222497; PubMed=10198195;
RA Lamb F.S., Clayton G.H., Liu B.-X., Smith R.L., Barna T.J.,
RA Schutte B.C.;
RT "Expression of CLCN voltage-gated chloride channel genes in human
RT blood vessels.";
RL J. Mol. Cell. Cardiol. 31:657-666(1999).
RN [4]
RP VARIANTS NPFL ARG-200; LEU-244; GLU-506 AND PRO-520.
RX MEDLINE=96158876; PubMed=8559248;
RA Lloyd S.E., Pearce S.H.S., Fisher S.E., Steinmeyer K., Schwappach B.,
RA Schelmann S.J., Harding B., Bolino A., Devoto M., Goodyer P.,
RA Rigden S.P.A., Wrong O., Jentsch T.J., Craig I.W., Thakker R.V.;
RT "A common molecular basis for three inherited kidney stone diseases.";
RL Nature 379:445-449(1996).
RN [5]
RP VARIANTS NPFL.
RX MEDLINE=97402204; PubMed=9259268;
RA Lloyd S.E., Guenther W., Pearce S.H.S., Thomson A., Bianchi M.L.,

RA Bosio M., Craig I.W., Fisher S.E., Scheinman S.J., Wrong O.,
 RA Jentsch T.J., Thakker R.V.;
 RT "Characterisation of renal chloride channel, CLCN5, mutations in
 RT hypercalcaemic nephrolithiasis (kidney stones) disorders.";
 RL Hum. Mol. Genet. 6:1233-1239(1997).
 CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
 CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
 CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
 CC TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL
 CC TUBULAR FUNCTION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: KIDNEY. MODERATELY EXPRESSED IN AORTIC
 CC VASCULAR SMOOTH MUSCLE AND ENDOTHELIAL CELLS, AND AT A SLIGHTLY
 CC HIGHER LEVEL IN THE CORONARY VASCULAR SMOOTH MUSCLE.
 CC -!- DISEASE: DEFECTS IN CLCN5 ARE THE CAUSE OF FOUR DISORDERS OF
 CC HEREDITARY HYPERCALCAEMIC NEPHROLITHIASIS (KIDNEY STONES), THAT
 CC HAVE BEEN REFERRED TO AS DENT'S DISEASE (DD), X-LINKED RECESSIVE
 CC NEPHROLITHIASIS (XRN), X-LINKED RECESSIVE HYPOPHOSPHATAEMIC
 CC RICKETS (XLRH) AND IDIOPATHIC LOW MOLECULAR WEIGHT PROTEINURIA OF
 CC JAPANESE CHILDREN (JILP). ALL FOUR DISEASES REPRESENT RENAL
 CC TUBULAR DISORDERS. THEY ARE CHARACTERIZED BY LOW MOLECULAR WEIGHT
 CC PROTEINURIA, HYPERCALCAEMIA, NEPHROCALCINOSIS, NEPHROLITHIASIS
 CC (KIDNEY STONES) AND RENAL FAILURE. DD IS A FORM OF FANCONI
 CC SYNDROME (ALSO KNOWN AS X-LINKED RECESSIVE NEPHROLITHIASIS TYPE 2
 CC (NPHL2)).
 CC -!- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91906; CAA63000.1; -;
 DR EMBL; X81836; CAA57430.1; -;
 DR Genew; HGNC:2023; CLCN5.
 DR MIM; 300008; -;
 DR MIM; 300009; -;
 DR MIM; 310468; -;
 DR InterPro; IPR000644; CBS_domain.
 DR InterPro; IPR001807; Cl-channel_volt.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 2.
 DR Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
 KW CBS domain; Repeat; Disease mutation.
 FT TRANSMEM 55 75 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT TRANSMEM 246 266 POTENTIAL.
 FT TRANSMEM 317 337 POTENTIAL.
 FT TRANSMEM 352 372 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT TRANSMEM 453 473 POTENTIAL.
 FT TRANSMEM 494 514 POTENTIAL.
 FT TRANSMEM 521 541 POTENTIAL.
 FT DOMAIN 585 645 CBS 1.
 FT DOMAIN 682 733 CBS 2.
 FT VARIANT 30 30
 FT R -> RH (IN DD).
 FT /FTid=VAR_001615.
 FT G -> V (IN DD).
 FT /FTid=VAR_001616.
 FT L -> R (IN NPHL2).
 FT /FTid=VAR_001617.
 FT S -> L (IN NPHL3).
 FT /FTid=VAR_001618.
 FT R -> P (IN JILP).
 FT /FTid=VAR_001619.
 FT G -> E (IN NPHL1).
 FT VARIANT 506 506

FT VARIANT 512 512 /FTid=VAR_001620.
 FT G -> R (IN DD; ABOLISHES THE CHLORIDE
 FT CURRENTS).
 FT /FTid=VAR_001621.
 FT S -> P (IN NPHL2).
 FT /FTid=VAR_001622.
 FT E -> D (IN DD; ABOLISHES THE CHLORIDE
 FT CURRENTS AND TOTAL LOSS OF FUNCTION).
 FT /FTid=VAR_001623.
 FT SEQUENCE 746 AA; 83146 MW; EF913C5BA40C85D8 CRC64;
 Query Match 58.9%; Score 43; DB 1; Length 746;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 CCGXXLCH 11
 Db 220 CCGNLTCH 227
 RESULT 8
 CLCN5_MOUSE
 ID CLCN5_MOUSE STANDARD; PRT; 746 AA.
 AC Q9WVD4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chloride channel protein 5 (CLC-5).
 GN CLCN5 OR CLCN5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=99303559; Pubmed=10373326;
 RA Tanaka K., Fisher S.E., Craig I.W.;
 RT "Characterization of novel promoter and enhancer elements of the mouse
 RT homologue of the dent disease gene, CLCN5, implicated in x-linked
 RT hereditary nephrolithiasis.";
 RL Genomics 58:281-292(1999).
 CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
 CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
 CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
 CC TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL
 CC TUBULAR FUNCTION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
 CC -!- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
 CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF134117; AAD28473.1; -;
 DR MGD; MGI:99486; Clcn5.
 DR InterPro; IPR000644; CBS_domain.
 DR InterPro; IPR001807; Cl-channel_volt.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 2.
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
 KW CBS domain; Repeat. 75
 FT TRANSMEM 55 75 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.


```

FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT DOMAIN 585 614 CBS 1.
FT DOMAIN 682 733 CBS 2.
SQ SEQUENCE 746 AA; 83100 MW; D8F3AE4FFC331A08 CRC64;

Query Match 58.9%; Score 43; DB 1; Length 746;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
   ||| |||
Db 220 CCGNILCH 227

RESULT 9
CLC5_RAT STANDARD; PRT; 746 AA.
ID CLC5_RAT STANDARD; PRT; 746 AA.
AC P51796; P70642; Rel. 34, Created
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLCN5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96125100; PubMed=8537381;
RA Steinmeyer K., Schwappach B., Bens M., Vandewalle A., Jentsch T.J.;
RT "Cloning and functional expression of rat CLC-5, a chloride channel
RT related to kidney disease".
RL J. Biol. Chem. 270:31172-31177(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96215316; PubMed=8626585;
RA Sakamoto H., Kawasaki M., Uchida S., Sasaki S., Marumo F.;
RT "Identification of a new outwardly rectifying Cl- channel that
RT belongs to a subfamily of the ClC Cl- channels."
RL J. Biol. Chem. 271:10210-10216(1996).
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL
CC TUBULAR FUNCTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
CC -!- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC -!- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 256277; CAA01216.1; -.
CC EMBL; D50497; BAA09091.1; -.
CC InterPro: IPR000644; CBS_domain.
CC InterPro: IPR001807; Cl-channel_volt.
CC Pfam; PF00571; CBS; 2.
CC Pfam; PF00654; voltage_CLC; 1.

```

```

DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW CBS domain; Repeat.
FT TRANSMEM 55 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT DOMAIN 585 614 CBS 1.
FT DOMAIN 682 733 CBS 2.
FT CONFLICT 315 - 315 H -> Y (IN REF. 2).
SQ SEQUENCE 746 AA; 83067 MW; 5F17D45F397003CE CRC64;

Query Match 58.9%; Score 43; DB 1; Length 746;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
   ||| |||
Db 220 CCGNILCH 227

RESULT 10
PGPA_ECOLI STANDARD; PRT; 172 AA.
ID PGPA_ECOLI STANDARD; PRT; 172 AA.
AC P18200; P77321;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphatidylglycerophosphatase A (pc 3.1.3.27).
GN PGPA OR B0418.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89033892; PubMed=2846510;
RA Icho T.;
RT "Membrane-bound phosphatases in Escherichia coli: sequence of the
RT pgpa gene."
RL J. Bacteriol. 170:5110-5116(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Nemati A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Iida A., Hayashi M., Fujio T., Teshiba S.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ONE OF THE THREE PHOSPHOLIPID PHOSPHATASES, SPECIFICALLY
CC HYDROLYZES PHOSPHATIDYLGLYCEROPHOSPHATE.
CC -!- CATALYTIC ACTIVITY: Phosphatidylglycerophosphate + H(2)O =
CC phosphatidylglycerol + phosphate.

```


DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Vasotocin-neurophysin VT precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White leghorn; TISSUE=Hypothalamus;
 RA Hunt N., Kluver D., Ivell R.;
 RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: VASOTOCIN IS A ANTIIDIURETIC HORMONE.
 CC -!- DOMAIN: IN NON-MAMMALIAN TETRAPODS, THE PROTEOLYTIC PROCESSING OF
 CC THE PRO-VASOTOCIN INVOLVES ONLY ONE CLEAVAGE, RELEASING THE
 CC HORMONE MOIEY AND A "BIG" NEUROPHYSIN WITH TWO DOMAINS HOMOLOGOUS
 CC TO THE MAMMALIAN NEUROPHYSIN II AND COPEPTIN, RESPECTIVELY.
 CC -!- PTM: SEVEN DISULFIDE BONDS ARE PRESENT IN NEUROPHYSIN.
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X55130; CAA38923.1; -
 DR PIR; S14480; S14480.
 DR HSSP; P01180; INPO.
 DR InterPro; IPR000981; Neurohyp_horm.
 DR Pfam; PF00184; hormone5; 1.
 DR Pfam; PF00220; hormone4; 1.
 DR PRINTS; PR00831; NEUROPHYSIN.
 DR ProDom; PD001676; Neurohyp_horm; 1.
 DR SMART; SM00003; NH; 1.
 DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
 KW Hormone; Hypothalamus; Cleavage on pair of basic residues;
 KW Amidation; Signal.
 FT SIGNAL 1 19
 FT PEPTIDE 20 28 VASOTOCIN.
 FT PEPTIDE 32 161 VT NEUROPHYSIN.
 FT DISULFID 20 25 BY SIMILARITY.
 FT MOD_RES 28 28 AMIDATION (G-29 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 161 AA; 16693 MW; 2802FBBED5E52277 CRC64;

 Query Match 56.2%; Score 41; DB 1; Length 161;
 Best Local Similarity 60.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

 QY 1 NGVCCGXLC 10
 DB 101 NGVCCSADTC 110

 RESULT 13
 ARDL_RAT
 ID ARDL_RAT STANDARD; PRT; 554 AA.
 AC P36407;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GTP-binding protein ARD-1 (Fragment).
 GN ARFD1 OR ARD1 OR ARD-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RP [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93232038; PubMed=8473324;
 RA

RA Mishima K., Tsuchiya M., Nightingale M.S., Moss J., Vaughan M.;
 RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-
 terminal ADP-ribosylation factor domain.";
 RL J. Biol. Chem. 268:8801-8807(1993).
 CC -!- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
 CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT. BELONGS TO THE ARF FAMILY
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION;
 CC OF GTP-BINDING PROTEINS.
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L04760; AAA41301.1; -
 DR HSSP; P32889; IRRG.
 DR InterPro; IPR000251; ARF family.
 DR InterPro; IPR003649; Bbox_C.
 DR InterPro; IPR005325; Small_GTP.
 DR InterPro; IPR000315; Znf_Bbox.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00025; arf; 1.
 DR Pfam; PF00643; zf-B_box; 1.
 DR SMART; SM00177; ARF; 1.
 DR SMART; SM00502; BBC; 1.
 DR SMART; SM00336; BBOX; 2.
 DR SMART; SM00184; RING; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS01019; ARF; 1.
 DR PROSITE; PS00119; ZF_BBOX; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW GTP-binding; Zinc-finger.
 FT NON_TER 1 1
 FT ZN_FING 11 56 RING-TYPE.
 FT ZN_FING 102 148 B BOX-TYPE.
 FT DOMAIN 370 554 ARF-LIKE.
 FT NP_BIND 391 398 GTP (BY SIMILARITY).
 FT NP_BIND 434 438 GTP (BY SIMILARITY).
 FT NP_BIND 493 496 GTP (BY SIMILARITY).
 SQ SEQUENCE 554 AA; 62187 MW; FB427D6F27680839 CRC64;

 Query Match 56.2%; Score 41; DB 1; Length 554;
 Best Local Similarity 28.6%; Pred. No. 34;
 Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

 QY 2 GVC-----CGXXLCHXC 13
 DB 12 GVCEDVFSLQGDVKVPRLLLCGHTVCHDC 39

 RESULT 14
 ARDL_HUMAN
 ID ARDL_HUMAN STANDARD; PRT; 574 AA.
 AC P36406; Q9BZY5; Q9BZY4;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GTP-binding protein ARD-1 (Tripartite motif protein 23).
 GN ARFD1 OR ARD1 OR TRIM23.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RP [1]
 RN SEQUENCE FROM N.A. (ISOFORM ALPHA).
 RP MEDLINE=93232038; PubMed=8473324;
 RA Mishima K., Tsuchiya M., Nightingale M.S., Moss J., Vaughan M.;

RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-
terminal ADP-ribosylation factor domain.";
RL J. Biol. Chem. 268:8801-8807(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).
RX MEDLINE=21231161; PubMed=11331580;
RA Raymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzi L.,
RA Riganelli D., Zanaria E., Messali S., Cainarca S., Guffanti A.,
RA Minucci S., Pellicci P.G., Ballabio A.;
RT "The tripartite motif family identifies cell compartments.";
RL EMBO J. 20:2140-2151(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NOT KNOWN, THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC -!- ALTERNATIVE PRODUCTS: 3 isoforms; Alpha (shown here), Beta and
CC Gamma; are produced by alternative splicing.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ARF FAMILY
CC OF GTP-BINDING PROTEINS.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L04510; AAA35940.1; -
DR EMBL; AF230397; AAG50176.1; -
DR EMBL; AF230398; AAG50177.1; -
DR EMBL; AF230399; AAG50178.1; -
DR EMBL; BC022510; AAH22510.1; -
DR PIR; A46054; A46054.
DR HSP; P32889; IRRG.
DR Genew; HGNC:660; ARD1.
DR MIM; 601747; -
DR InterPro; IPR000251; ARF_family.
DR InterPro; IPR003649; Bbox_C.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR000315; Znf_Bbox.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00025; arf; 1.
DR Pfam; PF00643; zf-B_box; 1.
DR SMART; SM00177; ARF; 1.
DR SMART; SM00502; BBC; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00184; RING; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS01019; ARF; 1.
DR PROSITE; PS01019; ZF_BBOX; 1.
DR PROSITE; PS00518; ZF_RING.1; 1.
DR PROSITE; PS00089; ZF_RING.2; 1.
KW GTP-binding; zinc-finger; Alternative splicing.
FT ZN_FING 31 76
FT ZN_FING 122 168 RING-TYPE.
FT DOMAIN 390 574 B BOX-TYPE.
FT NP_BIND 411 418 GTP (BY SIMILARITY).
FT NP_BIND 454 458 GTP (BY SIMILARITY).
FT NP_BIND 513 516 GTP (BY SIMILARITY).
FT VARSPIC 551 574 GMGLVEGLDWLSRQLVAAGVLDVA -> VFQICDQYTGKE
FT VARSPLIC 541 574 VVTEKG (IN ISOFORM BETA).
FT VARSPLIC 574 574 WYIGCCDARSNGLYEGLDWLSRQLVAAGVLDVA -> CFS
FT DNM (IN ISOFORM GAMMA).
SQ SEQUENCE 574 AA; 64066 MW; CB85923B29BF0320 CRC64;
Query Match 56.2%; Score 41; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 35;

Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;
QY 2 GVC-----CGXXLCHXC 13
DB 32 GVCEDVFSLOGDKVPRLLLCGHTVCHDC 59
||| ||:| |
||| |
RESULT 15
UBPA_DICDI
ID UBPA_DICDI STANDARD; PRT; 837 AA.
AC P54201;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase A (EC 3.1.2.15) (Ubiquitin
DE thiolesterase A) (Ubiquitin-specific processing protease A)
DE (Deubiquitinating enzyme A).
GN UBPA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99003275; PubMed=9786928;
RA Lindsey D.F., Amerik A., Deery W.J., Bishop J.D., Hochstrasser M.,
RA Gomer R.H.;
RT "A deubiquitinating enzyme that disassembles free polyubiquitin
RT chains is required for development but not growth in Dictyostelium.";
RL J. Biol. Chem. 273:29178-29187(1998).
CC -!- FUNCTION: REQUIRED FOR DEVELOPMENT BUT NOT GROWTH.
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U48271; AAC71068.1; -
DR MEROPS; C19.0PW; -
DR DictyDb; DD00088; ubpa.
DR InterPro; IPR000449; USA_domain.
DR InterPro; IPR001394; UCH-2.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00442; UCH-1; 1.
DR Pfam; PF00443; UCH-2; 1.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF02148; zf-UBP; 1.
DR SMART; SM00185; UBA; 2.
DR SMART; SM00290; Znf_UBP; 1.
DR PROSITE; PS00972; UCH_2.1; 1.
DR PROSITE; PS00973; UCH_2.2; 1.
DR PROSITE; PS50235; UCH_2.3; 1.
KW Ubl conjugation pathway; Hydrolase; Thiol protease.
FT ACT_SITE 328 328 BY SIMILARITY.
FT ACT_SITE 788 788 BY SIMILARITY.
FT ACT_SITE 797 797 BY SIMILARITY.
FT DOMAIN 684 690 POLY-ASN.
SQ SEQUENCE 837 AA; 94978 MW; 98282776436E52D3 CRC64;
Query Match 54.8%; Score 40; DB 1; Length 837;
Best Local Similarity 54.5%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 NGVCGGXXLCH 11
DB 791 NNVTGCHVCH 801

Search completed: January 29, 2003, 09:29:23
Job time : 3.95455 secs

THIS PAGE BLANK (user)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:57 ; Search time 5.46591 Seconds
(without alignments)
228.644 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCCGXXLCHXC 13
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	63.0	258	T30368	probable immediate
2	46	63.0	1322	T15689	hypothetical prote
3	45	61.6	400	ZBBE14	44.1K zinc-binding
4	44	60.3	1235	1 QBBE4	DNA-binding protei
5	43	58.9	104	2 F83370	hydrogen cyanide s
6	42	57.5	171	2 G90687	phosphatidylglycer
7	42	57.5	171	2 C85538	phosphatidylglycer
8	42	57.5	171	2 A00554	phosphatidylglycer
9	42	57.5	172	2 B64771	phosphatidylglycer
10	42	57.5	244	2 T6913	hypothetical prote
11	42	57.5	322	2 T6912	hypothetical prote
12	42	57.5	322	2 T08592	TGACG-motif-bindin
13	42	57.5	326	2 T08591	TGACG-motif bindin
14	42	57.5	332	2 T24312	hypothetical prote
15	42	57.5	1184	2 T00253	gene Anknzn protei
16	41.5	56.8	768	2 B41029	integrin beta-8 ch
17	41	56.2	161	2 S14480	arginine-vasotocin
18	41	56.2	179	2 A83068	hypothetical prote
19	41	56.2	328	2 E97129	probable flavodoxi
20	41	56.2	514	2 T10559	hypothetical prote
21	41	56.2	557	2 T27752	hypothetical prote
22	41	56.2	574	2 A46054	GTP-binding protei
23	40	54.8	98	2 E81205	hypothetical prote
24	40	54.8	122	2 D90867	hypothetical prote
25	40	54.8	122	2 E85751	hypothetical prote
26	40	54.8	550	2 G85436	hypothetical prote
27	40	54.8	1101	2 T16840	hypothetical prote
28	39	53.4	55	2 S25774	testis-specific pr
29	39	53.4	144	2 A29101	vasopressin / neur

30	39	53.4	149	2	T21037	hypothetical prote
31	39	53.4	186	2	A45910	ultra-high-sulfur
32	39	53.4	211	2	T20590	hypothetical prote
33	39	53.4	245	2	T10407	immediate early pr
34	39	53.4	286	2	S61199	hypothetical prote
35	39	53.4	303	2	A87616	hydroxylase, carbon-
36	39	53.4	1160	2	A36256	DNA-binding protei
37	38.5	52.7	3175	1	RRWVEV	genome polyprotein
38	38	52.1	91	2	S02769	gag 75K protein pr
39	38	52.1	98	2	JC5147	tachycitin precurs
40	38	52.1	155	2	B32669	vasotocin 2 / neur
41	38	52.1	250	2	T29344	hypothetical prote
42	38	52.1	351	2	S76925	hypothetical prote
43	38	52.1	553	2	B88072	protein ZK1240.2 l
44	38	52.1	658	2	T33568	hypothetical prote
45	38	52.1	948	2	A57640	retinoblastoma bin

ALIGNMENTS

RESULT 1

T30368
probable immediate-early transactivator 0 - Lymantria dispar nuclear polyhedrosis vir
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30368
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A:Reference number: Z20836; MUID:99124785; PMID:9887315
A:Accession: T30368
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-258 <KUZ>
A:Cross-references: EMBL:AF081810; PIDN:AAC70206.1
C:Keywords: immediate-early protein

Query Match 63.0%; Score 46; DB 2; Length 258;
Best Local Similarity 54.5%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VCCGXXLCHXC 13
|||||:|:
Db 208 VCCGYRVCNAC 218

RESULT 2

T15689
hypothetical protein C28G1.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C:Accession: T15689
R:Favella, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C28G1.
A:Reference number: Z18389
A:Accession: T15689
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1322 <FAV>
A:Cross-references: EMBL:U41026; NID:g1086701; PID:g1086702; PIDN:AAA82350.1; CESP:C2
C:Genetics:
A:Gene: CESP:C28G1.3
A:Introns: 25/; 131/3; 150/1; 166/3; 180/3; 204/2; 235/3; 344/3; 385/2; 436/3; 483/2
C:Superfamily: RING finger homology
F:810-862/Domain: RING finger homology <RRN>

Query Match 63.0%; Score 46; DB 2; Length 1322;
Best Local Similarity 63.6%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VCCGXXLCHXC 13

Db 1071 VTGHALCHKC 1081

RESULT 3

ZBBE14

44.1K zinc-binding protein - ictalulrid herpesvirus 1 (strain auburn 1)

C:Species: ictalulrid herpesvirus 1

A:Note: host Ictalurus punctatus (channel catfish)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: D36794

R:Davidson, A.J.

submitted to GenBank, January 1992

A:Description: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A36804

A:Accession: D36794

A:Molecule type: DNA

A:Residues: 1-400 <DAV>

A:Cross-references: GB:M75136; NID:g331209; PIDN:AAA88180.1; PID:g331286

R:Davidson, A.J.

Virology 186, 9-14, 1992

A:Title: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A39447; MUID:92087490; PMID:1727613

A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 78

C:Superfamily: ictalulrid herpesvirus 44.1K zinc binding protein

C:Keywords: zinc finger

Query Match

Best Local Similarity 61.6%; Score 45; DB 1; Length 400;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 CCGXXLCHXC 13

Db 133 CCGATLCDSC 142

RESULT 4

QDBEWA

DNA-binding protein - human cytomegalovirus (strain AD169)

N:Alternate names: UL57 protein

C:Species: human cytomegalovirus, human herpesvirus 5

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: S09820

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;

M.; Barré, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; MUID:90269039; PMID:2161319

A:Accession: S09820

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1235 <CHE>

A:Cross-references: EMBL:X17403; NID:q59591; PIDN:CAA35372.1; PID:g1780835

A:Note: possible protein-coding frames are given

A:Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form

C:Superfamily: herpesvirus DNA-binding protein

C:Keywords: DNA binding

Query Match

Best Local Similarity 60.3%; Score 44; DB 1; Length 1235;

Matches 7; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 GVC--CGXXLCHXC 13

Db 470 GLCEAGGTCCHTC 483

RESULT 5

F83370

hydrogen cyanide synthase HcnA PA2193 [imported] - Pseudomonas aeruginosa (strain PAO

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: F83370

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: F83370

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-104 <STO>

A:Cross-references: GB:AE004646; GB:AE004091; NID:g9948213; PIDN:AAG05581.1; GSPDB:GN

A:Experimental source: strain PAO1

C:Genetics:

A:Gene: hcnA; PA2193

Query Match

Best Local Similarity 58.9%; Score 43; DB 2; Length 104;

Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13

Db 56 SGAFCGMGVCHCC 68

RESULT 6

G90687

phosphatidylglycerophosphatase [imported] - Escherichia coli (strain O157:H7, substra

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: G90687

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: G90687

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-171 <HAY>

A:Cross-references: PIDN:BA000007; PIDN:BA033894.1; PID:q13359928; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: ECS0471

C:Superfamily: conserved hypothetical protein H11306

Query Match

Best Local Similarity 57.5%; Score 42; DB 2; Length 171;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

Db 67 GICIGWLCH 76

RESULT 7

C85538

phosphatidylglycerophosphatase [imported] - Escherichia coli (strain O157:H7, substra

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: C85538

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: C85538

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-171 <STO>

A:Cross-references: GB:AB005174; NID:gl2513273; PIDN:AA654767.1; GSPDB:GN00145; UWGP:Z05
A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: pppA

C:Superfamily: conserved hypothetical protein H11306

Query Match 57.5%; Score 42; DB 2; Length 171;
Best Local Similarity 60.0%; Pred. NO. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

I:| | | | |

Db 67 GICIGVYLCH 76

RESULT 8

AD0554

phosphatidylglycerophosphatase A [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB0554

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
J. S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0554

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-171 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08876.1; PID:gl6501688; GSPDB:GN00176

C:Genetics:

A:Gene: STY0459

C:Superfamily: conserved hypothetical protein H11306

Query Match 57.5%; Score 42; DB 2; Length 171;
Best Local Similarity 60.0%; Pred. NO. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

I:| | | | |

Db 67 GICIGVYLCH 76

RESULT 9

B64771

phosphatidylglycerophosphatase (EC 3.1.3.27) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: B64771

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: B64771

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-172 <BLAT>

A:Cross-references: GB:AB000148; GB:U00096; NID:gl786614; PIDN:AA673521.1; PID:gl786620;

A:Experimental source: strain K-12, substrain MGL655

C:Genetics:

A:Gene: pppA

C:Function:

A:Pathway: phospholipid degradation

C:Superfamily: conserved hypothetical protein H11306

C:Keywords: phosphoric monoester hydrolase; transmembrane protein

F:59-75/Domain: transmembrane #status predicted <TM1>

F:143-159/Domain: transmembrane #status predicted <TM2>

Query Match 57.5%; Score 42; DB 2; Length 172;
Best Local Similarity 60.0%; Pred. NO. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

I:| | | | |

Db 67 GICIGVYLCH 76

RESULT 10

T26913

hypothetical protein Y45F10B.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26913

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20286

A:Accession: T26913

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-244 <WIL>

A:Cross-references: EMBL:AL021487; PIDN:CAA16351.1; GSPDB:GN00022; CESP:Y45F10B.8

A:Experimental source: clone Y45F10B

C:Genetics:

A:Gene: CESP:Y45F10B.8

A:Map position: 4

A:Introns: 90/2; 194/1

Query Match 57.5%; Score 42; DB 2; Length 244;
Best Local Similarity 55.6%; Pred. NO. 37;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 CGXXLCHXC 13

I:| | | | |

Db 195 CGHTTCHTC 203

RESULT 11

T26912

hypothetical protein Y45F10B.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26912

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20286

A:Accession: T26912

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-244 <WIL>

A:Cross-references: EMBL:AL021487; PIDN:CAA16350.1; GSPDB:GN00022; CESP:Y45F10B.9

A:Experimental source: clone Y45F10B

C:Genetics:

A:Gene: CESP:Y45F10B.9

A:Map position: 4

A:Introns: 90/2; 194/1

Query Match 57.5%; Score 42; DB 2; Length 244;
Best Local Similarity 55.6%; Pred. NO. 37;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 CGXXLCHXC 13

I:| | | | |

Db 195 CGHTTCHTC 203

RESULT 12

T08592

TCACG-motif-binding protein STF2 - soybean

C:Species: Glycine max (soybean)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999

C:Accession: T08592

R;Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Goekjian, V.H.; Nagao, R.T.; Key, J.L.
 submitted to the EMBL Data Library, September 1995
 A:Description: STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bZIP
 A:Reference number: Z16445
 A:Accession: T08592
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-322 <CHE>
 A:Cross-references: EMBL:L28004; NID:g986966; PID:g2934885
 A:Experimental source: strain Williams; hypocotyl
 C:Genetics:
 A:Gene: STF2

Query Match 57.5%; Score 42; DB 2; Length 322;
 Best Local Similarity 66.7%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
 || ||| |
 Db 61 CGFPLCHSC 69

RESULT 13
 T08591
 TGACG-motif binding protein STF1 - soybean
 C:Species: Glycine max (soybean)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
 C:Accession: T08591
 R;Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Goekjian, V.H.; Nagao, R.T.; Key, J.L.
 submitted to the EMBL Data Library, September 1995
 A:Description: STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bZIP
 A:Reference number: Z16445
 A:Accession: T08591
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-326 <CHE>
 A:Cross-references: EMBL:L28003; NID:g2934883; PID:g2934884
 A:Experimental source: strain Williams; hypocotyl

Query Match 57.5%; Score 42; DB 2; Length 326;
 Best Local Similarity 66.7%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
 || ||| |
 Db 61 CGFPLCHSC 69

RESULT 14
 T24312
 hypothetical protein T01G5.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T24312
 R;Basham, V.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19873
 A:Accession: T24312
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-332 <SWIL>
 A:Cross-references: EMBL:Z81111; PIDN:CAB03269.1; GSPDB:GN00023; CESP:T01G5.7
 A:Experimental source: clone T01G5
 C:Genetics:
 A:Gene: CESP:T01G5.7
 A:Map position: 5
 A:Introns: 208/1; 257/1

Query Match 57.5%; Score 42; DB 2; Length 332;
 Best Local Similarity 55.6%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13

Db 209 CGHTMCHTC 217
 || :|| |
 RESULT 15
 T00253
 gene Ankhzn protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 05-Nov-1999
 C:Accession: T00253
 R:Ito, K.; Miyashita, A.; Ishii, N.; Kuriyama, H.; Tominaga, K.; Saitoh, H.; Maruyama
 submitted to the EMBL Data Library, February 1998
 A:Description: CDNA cloning of mouse Ankhzn.
 A:Reference number: Z14132
 A:Accession: T00253
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-1184 <ITO>
 A:Cross-references: EMBL:AB011370; NID:d1179960; PIDN:BAA24980.1; PID:d1025905
 C:Genetics:
 A:Note: Ankhzn

Query Match 57.5%; Score 42; DB 2; Length 1184;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
 || ||| |
 Db 1144 CGRLCHKC 1152

Search completed: January 29, 2003, 09:32:59
 Job time : 6.46591 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2003, 09:28:52 ; Search time 13 Seconds
(without alignments)
133.251 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	90.4	13	21 AAB08013	Generic formula fo
2	66	90.4	13	21 AAB08017	Amino acid sequenc
3	66	90.4	13	21 AAY92229	Chi-conotoxin pept
4	66	90.4	61	21 AAB08016	Amino acid sequenc
5	66	90.4	61	21 AAY92231	Chi-conotoxin, chi
6	60	82.2	12	21 AAB08014	Generic formula fo
7	60	82.2	12	21 AAB08015	Generic formula fo
8	60	82.2	12	21 AAB08018	Amino acid sequenc
9	60	82.2	12	21 AAB08019	Amino acid sequenc
10	60	82.2	13	21 AAY92230	Chi-conotoxin pept

11	44	60.3	538	22	AAU06111	Novel human polype
12	44	60.3	538	23	AAU97060	Human Ajuba-like p
13	43	58.9	93	22	AAU31406	Novel human secret
14	43	58.9	766	22	AAM79259	Human protein SFQ
15	43	58.9	1203	22	ABG05471	Novel human diagno
16	43	58.9	1329	23	AAU91279	Human NOV3a protei
17	43	58.9	1597	22	ABG03989	Novel human diagno
18	43	58.9	1597	22	ABG10253	Novel human diagno
19	43	58.9	1597	22	ABG10812	Novel human diagno
20	43	58.9	1597	22	ABG11902	Novel human diagno
21	43	58.9	1599	22	ABG07884	Novel human diagno
22	43	58.9	1599	22	ABG14450	Novel human diagno
23	43	58.9	1784	22	ABG09148	Novel human diagno
24	42	57.5	369	22	ABG15487	Novel human diagno
25	42	57.5	865	22	ABG25368	Novel human diagno
26	42	57.5	1129	22	AAAB62174	Human SV protein.
27	42	57.5	1169	22	AAAB62173	Human Rabankyrin-5
28	42	57.5	1170	22	AAW40022	Human polypeptide
29	42	57.5	1314	23	ABB92290	Herbicide beta-8 inte
30	41.5	56.8	768	13	AAAB08012	Generic formula fo
31	41	56.2	13	21	AAAB08012	Human secreted pro
32	41	56.2	60	20	AAAY27648	Human secreted pro
33	41	56.2	164	21	AAAG01721	Propionibacterium
34	41	56.2	360	22	AAU64532	Human parkin gene
35	41	56.2	437	20	AAAY32502	Human parkin gene
36	41	56.2	465	20	AAAY32501	Human parkin gene
37	41	56.2	574	15	AAAB66034	Rat ARD 1. Rattus
38	41	56.2	574	15	AAAB66033	Human ARD 1. Homo
39	41	56.2	919	22	AAAG84970	Shrimp white spot
40	40	54.8	73	19	AAAY20881	Human presenilin I
41	40	54.8	136	23	ABP04912	Human OREF protein
42	40	54.8	344	22	AAAB67521	Amino acid sequenc
43	40	54.8	451	22	AAAB67533	Amino acid sequenc
44	40	54.8	464	22	AAAB67517	Amino acid sequenc
45	40	54.8	464	22	AAAB67531	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAB08013
ID AAB08013 standard; peptide; 13 AA.
XX AC AAB08013;
XX DT 14-NOV-2000 (first entry)
XX DE Generic formula for conotoxin peptide Marl.
XX KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
XX OS Conus marmoreus.
XX FH Key Location/Qualifiers
FT Misc-difference 7 /label= Tyr, Xaa
FT FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT FT Misc-difference 8 /label= Lys, Xaa
FT FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT FT N,N,N-trimethyl Lys"
FT FT Misc-difference 12 /note= "optionally hydroxy-Pro"
FT FT WO200044769-A1.
XX PN
XX PD 03-AUG-2000.
XX PF 28-JAN-2000; 2000WO-US01978.
XX PR 29-JAN-1999; 99US-0118381.

PR 28-DEC-1999; 99US-0173343.
 XX (UTAH) UNIV UTAH RES FOUND.
 PA
 XI McIntosh JM, Olivera BM, Cruz LJ;
 XX WPI; 2000-476222/41.
 XX Purified ap-conotoxin derived from cone snail venom for use as an
 PT analgesic -
 XX Claim 2; Page 19; 29pp; English.
 XX The present sequence represents an ap-conotoxin peptide, designated
 CC Marl. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 XX SQ Sequence 13 AA;
 Query Match 90.4%; Score 66; DB 21; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.017; Mismatches 0; Gaps 0;
 Matches 12; Conservative 0; Indels 1; Indels 0; Gaps 0;
 QY 1 NGVCCGXXLCHXC 13
 DB 1 NGVCCGXXLCHPC 13
 RESULT 2
 AAB08017
 ID AAB08017 standard; peptide; 13 AA.
 AC AAB08017;
 DT 14-NOV-2000 (first entry)
 DE Amino acid sequence of the conotoxin peptide Marl.
 KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
 OS Conus marmoreus.
 XX Key Location/Qualifiers
 FH Modified-site 12 /note= "hydroxy-Pro"
 FT
 XX WO200044769-A1.
 XX 03-AUG-2000.
 XX 28-JAN-2000; 2000WO-US01978.
 XX 29-JAN-1999; 99US-0118381.
 PR 28-DEC-1999; 99US-0173343.
 XX (UTAH) UNIV UTAH RES FOUND.
 PA
 XI McIntosh JM, Olivera BM, Cruz LJ;
 XX WPI; 2000-476222/41.
 XX Purified ap-conotoxin derived from cone snail venom for use as an
 PT analgesic -
 XX Claim 13; Page -; 29pp; English.
 XX The present sequence represents an ap-conotoxin peptide, designated
 CC Marl. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.

XX SQ Sequence 13 AA;
 Query Match 90.4%; Score 66; DB 21; Length 13;
 Best Local Similarity 76.9%; Pred. No. 0.017; Mismatches 0; Gaps 0;
 Matches 10; Conservative 0; Indels 3; Indels 0; Gaps 0;
 QY 1 NGVCCGXXLCHXC 13
 DB 1 NGVCCGXXLCHPC 13
 RESULT 3
 AAY92229
 ID AAY92229 standard; peptide; 13 AA.
 XX AAY92229;
 AC AAY92229;
 DT 10-AUG-2000 (first entry)
 DE Chi-conotoxin peptide, chi-MrIA.
 KW chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenaline transporter; urinary tract disorder; analgesic;
 KW antiarrhythmic; cardiant; antidepressant; anxiolytic; anti-inflammatory.
 OS Conus marmoreus.
 XX Key Location/Qualifiers
 FH Misc-difference 12 /label= 4Hyp
 FT /note= "4-hydroxyproline"
 XX WO200020444-A1.
 XX 13-APR-2000.
 XX 01-OCT-1999; 99WO-AU00844.
 XX 02-OCT-1998; 98AU-0006274.
 XX (UYQU) UNIV QUEENSLAND.
 XX Lewis RJ, Alewood PF, Sharpe IA;
 DR WPI; 2000-303738/26.
 XX Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation
 XX Claim 3; Page 33; 47pp; English.
 XX This conotoxin, chi-MrIA, is a member of a new class of conotoxins,
 CC designated chi-conotoxin. It was isolated from the venom of the mollusc
 CC hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
 CC neuronal amine transporters, especially the neuronal noradrenaline
 CC transporter. Inhibitors of noradrenaline re-uptake which have a
 CC negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chi-MrIA (0.1 nM-1 micro M)
 CC inhibited the accumulation of radiolabeled noradrenaline in a
 CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are
 CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).
 XX

SQ Sequence 13 AA;
 Query Match 90.4%; Score 66; DB 21; Length 13;
 Best Local Similarity 76.9%; Pred. No. 0.017;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
 ||||| ||| I
 Db 1 NGVCCGYKLCHPC 13

RESULT 4
 AAB08016
 ID AAB08016 standard; Protein; 61 AA.
 AC AAB08016;
 XX
 DT 14-NOV-2000 (first entry)
 XX
 DE Amino acid sequence of the conotoxin Marl propeptide.
 XX
 KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
 XX
 OS Conus marmoreus.
 XX
 PN WO200044769-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-US01978.
 XX
 PR 29-JAN-1999; 99US-0118381.
 XX
 PR 28-DEC-1999; 99US-0173343.
 XX
 PA (UTAH) UNIV UTAH RES FOUND.
 XX
 PI McIntosh JM, Olivera BM, Cruz LJ;
 XX
 DR WPI; 2000-476222/41.
 XX
 DR N-PSDB; AAA63513.
 XX
 PT Purified ap-conotoxin derived from cone snail venom for use as an
 analgesic -
 XX
 PS Claim 23; Page 13-14; 29pp; English.
 XX
 CC The present sequence represents a Marl propeptide. Marl is an
 ap-conotoxin peptide. Conotoxins are naturally available in minute
 amounts in the venom of cone snails. The peptides have analgesic
 activity. The peptides are used to treat or prevent pain.
 XX
 SQ Sequence 61 AA;
 Query Match 90.4%; Score 66; DB 21; Length 61;
 Best Local Similarity 76.9%; Pred. No. 0.058;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
 ||||| ||| I
 Db 49 NGVCCGYKLCHPC 61

RESULT 5
 AAY92231
 ID AAY92231 standard; Protein; 61 AA.
 XX
 AC AAY92231;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Chi-conotoxin, chi-MrIA, leader and mature peptide.
 XX
 KW chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;

KW neuronal; noradrenaline transporter; urinary tract; analgesic; cardiant;
 antiarrhythmic; antidepressant; anxiolytic; anti-inflammatory.
 XX
 OS Conus marmoreus.
 XX
 PN WO200020444-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 01-OCT-1999; 99WO-AU00844.
 XX
 PR 02-OCT-1998; 98AU-0006274.
 XX
 PA (UYQU) UNIV QUEENSLAND.
 XX
 PI Lewis RJ, Alewood PF, Sharpe IA;
 XX
 DR WPI; 2000-303738/26.
 XX
 DR N-PSDB; AAA09112.
 XX
 PT Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 inhibiting neuronal amine transporter used for treatment or prophylaxis
 of urinary or cardiovascular conditions, mood disorders, or
 treatment/control of pain/inflammation
 XX
 PS Example 7; Page 31; 47pp; English.
 XX
 CC This sequence is the conotoxin, chi-MrIA, a member of a new class of
 conotoxins, designated chi-conotoxin. It was isolated from the venom of
 the mollusc hunting cone snail, Conus marmoreus. The peptide is an
 inhibitor of the neuronal amine transporters, especially the neuronal
 noradrenaline transporter. Inhibitors of noradrenaline re-uptake which
 have a negligible anti-cholinergic effect are particularly useful in the
 treatment of lower urinary tract disorders. Chi-MrIA (0.1 nM-1 micro M)
 inhibited the accumulation of radiolabeled noradrenaline in a
 concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 the accumulation by 50 percent was found to be approximately 7 nM. This
 concentration is approximately one order of magnitude lower than that
 needed for desipramine to produce the same effect. The peptides are
 useful for the treatment or prophylaxis of urinary or cardiovascular
 conditions or diseases (arrhythmia or coronary heart failure) or mood
 disorders (depression, anxiety or cravings), or the treatment or control
 of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 pain).
 XX
 SQ Sequence 61 AA;
 Query Match 90.4%; Score 66; DB 21; Length 61;
 Best Local Similarity 76.9%; Pred. No. 0.058;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
 ||||| ||| I
 Db 49 NGVCCGYKLCHPC 61

RESULT 6
 AAB08014
 ID AAB08014 standard; peptide; 12 AA.
 XX
 AC AAB08014;
 XX
 DT 14-NOV-2000 (first entry)
 XX
 DE Generic formula for conotoxin peptide Mar2.
 XX
 KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
 XX
 OS Conus marmoreus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 6

CC The present sequence represents an ap-conotoxin peptide, designated
 CC Mar2. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.

XX Sequence 12 AA;

Query Match 82.2%; Score 60; DB 21; Length 12;
 Best Local Similarity 75.0%; Pred. No. 0.094; Mismatches 3; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

QY 2 GVCCGXLCXHC 13
 IIII IIII
 Db 1 GVCCGYKLCHPC 12

RESULT 9
 AAB08019
 ID AAB08019 standard; peptide; 12 AA.

XX AAB08019;

DT 14-NOV-2000 (first entry)

DE Amino acid sequence of the conotoxin peptide U036.

KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.

XX Conus marmoreus.

FS Key Location/Qualifiers

FT Misc-difference 7

FT /label= Lys, Xaa

FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,

FT N,N'-trimethyl Lys"

FT Modified-site 11

FT /note= "hydroxy-Pro"

XX WO200044769-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US01978.

XX 29-JAN-1999; 99US-0118381.

PR 28-DEC-1999; 99US-0173343.

XX (UTAH) UNIV UTAH RES FOUND.

XX McIntosh JM, Olivera BM, Cruz LJ;

XX WPI; 2000-476222/41.

XX Purified ap-conotoxin derived from cone snail venom for use as an
 XX analgesic -

XX Claim 15; Page -: 29pp; English.

XX The present sequence represents an ap-conotoxin peptide, designated
 CC U036. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.

CC note: this sequence does not appear in the specification; it was created
 CC using information provided.

XX Sequence 12 AA;

Query Match 82.2%; Score 60; DB 21; Length 12;
 Best Local Similarity 83.3%; Pred. No. 0.094; Mismatches 2; Indels 0; Gaps 0;
 Matches 10; Conservative 0;

QY 2 GVCCGXLCXHC 13
 IIII IIII
 Db 1 GVCCGYKLCHPC 12

RESULT 10

AAAY92230

ID AAY92230 standard; peptide; 13 AA.

XX AAY92230;

DT 10-AUG-2000 (first entry)

XX Chi-conotoxin peptide, chi-MrIB.

KW chi-conotoxin; chi-MrIB; cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenaline transporter; urinary tract disorder; analgesic;
 KW antiarrhythmic; cardiant; antidepressant; anxiolytic; anti-inflammatory.

XX Conus marmoreus.

FS Key Location/Qualifiers

FT Misc-difference 12

FT /label= 4Hyp

FT /note= "4-hydroxyproline"

XX WO200020444-A1.

XX 13-APR-2000.

XX 01-OCT-1999; 99WO-AU00844.

XX 02-OCT-1998; 98AU-0006274.

XX (UYQU) UNIV QUEENSLAND.

XX Lewis RJ, Alewood PF, Sharpe IA;

XX WPI; 2000-303738/26.

XX Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation

XX Claim 3; Page 33; 47pp; English.

XX This conotoxin, chi-MrIB, is a member of a new class of conotoxins,
 CC designated chi-conotoxin. It was isolated from the venom of the mollusc
 CC hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
 CC neuronal amine transporters, especially the neuronal noradrenaline
 CC transporter. Inhibitors of noradrenaline re-uptake which have a
 CC negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chi-MrIA (0.1 nM-1 micro M)
 CC inhibited the accumulation of radiolabeled noradrenaline in a
 CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are
 CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).

XX Sequence 13 AA;

Query Match 82.2%; Score 60; DB 21; Length 13;
 Best Local Similarity 75.0%; Pred. No. 0.1;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXLCXHC 13

```

Db      2  GVCCGYKLCHPC 13
||||| ||| |
RESULT 11
AAU06111
ID  AAU06111 standard; Protein; 538 AA.
XX  AC  AAU06111;
XX  DT  24-OCT-2001 (first entry)
XX  DE  Novel human polypeptide PROT1.
XX  KW  Human; PROT1; AJUBA-like; central nervous system disorder; CNS;
XX  KW  cancer; neuromuscular disorder; cardiac disorder; clone AL132780A.
XX  OS  Homo sapiens.
XX  PN  WO200158946-A2.
XX  PD  16-AUG-2001.
XX  PF  08-FEB-2001; 2001WO-US04402.
XX  PR  08-FEB-2000; 2000US-0180880.
XX  PR  08-FEB-2000; 2000US-0181044.
XX  PR  10-FEB-2000; 2000US-0181656.
XX  PR  15-FEB-2000; 2000US-0182795.
XX  PR  07-FEB-2001; 2001US-0182795.
XX  PA  (CURA-) CURAGEN CORP.
XX  PI  Taupier RJ, Majumder K, Vernet CAM, Prayaga SK;
XX  WI  WPI; 2001-488970/53.
XX  N-PSDB; AAS09145.
XX  Isolated AJUBA-like, keratin-like and endothelin polypeptides for the
XX  treatment and diagnosis of cancers and neurological disorders -
XX  Claim 1; Page 9-10; 141pp; English.
XX  The present invention relates to the isolation of 4 novel human
XX  polypeptides, termed PROT1 polypeptides (AAU06111-AAU06114). The
XX  PROT1 polypeptides are AJUBA-like, keratin-like or endothelin-
XX  polypeptides. Polynucleotides encoding PROT1, PROT2 polypeptides and
XX  antibodies that bind PROT1 polypeptides are useful for treating or
XX  preventing a pathology associated with altered levels of PROT1 especially
XX  in a human subject. PROT1 polypeptides, PROT1 polynucleotides and PROT1
XX  antibodies are useful for identifying an agent that binds to PROT1, for
XX  identifying potential therapeutic agents for use in a treatment of a
XX  pathology related to aberrant expression or physiological interaction
XX  with PROT1, for screening for a modulator of activity of PROT1 and
XX  for determining the presence of or a predisposition to a disease
XX  associated with altered levels of PROT1. PROT1, an AJUBA-like protein,
XX  is useful for treating central nervous system disorders, cancer,
XX  Parkinson's disease, Alzheimer's disease, neuromuscular and cardiac
XX  disorders. PROT2, a keratin, is useful for treating inflammatory and
XX  neoplastic pancreatic disorders. PROT4, an endothelin, is useful for
XX  treating hypertension, cardiovascular and neurological disorders. The
XX  present sequence represents human polypeptide PROT1. The DNA sequence
XX  encoding PROT1 is isolated from the genomic clone AL132780_A.
XX  Sequence 538 AA;
XX  Query Match 60.3%; Score 44; DB 22; Length 538;
XX  Best Local Similarity 50.0%; Pred. No. 2.3e+02;
XX  Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY  2  GVCC-----GXXLCHXC 13
    | | | | |
Db  508  GCCCFPLDGHLLCHGC 523

RESULT 13
AAU31406
ID  AAU31406 standard; Protein; 93 AA.

```

```

RESULT 12
AAU97060
ID  AAU97060 standard; Protein; 538 AA.
XX  AC  AAU97060;
XX  DT  13-AUG-2002 (first entry)
XX  DE  Human Ajuba-like protein MFQ-114.
XX  KW  Human; Ajuba-like protein; MFQ-114; breast; prostate; ovary; pancreatic;
XX  KW  cancer; neoplastic disorder; neurodegenerative disease; stroke;
XX  KW  cardiac disease; vascular disease; angiogenesis; vaccine; immunogen;
XX  KW  gene therapy.
XX  OS  Homo sapiens.
XX  PN  WO200234911-A1.
XX  PD  02-MAY-2002.
XX  PF  20-SEP-2001; 2001WO-EP10882.
XX  PR  20-OCT-2000; 2000EP-0122857.
XX  PA  (MERE ) MERCK PATENT GMBH.
XX  PI  Soler Riera M, Masa Alvarez M, Hernandez Miguez JL, Piulats Xanco J;
XX  PI  Rosell Vives E;
XX  WI  WPI; 2002-463358/49.
XX  N-PSDB; ABK52759.
XX  Novel human Ajuba like protein, designated MFQ-114 polypeptide, useful
XX  for treating cancer, neurodegenerative disease, brain stroke, cardiac
XX  and vascular disease, and angiogenesis -
XX  Claim 1; Page 40-41; 45pp; English.
XX  The invention relates to an isolated human Ajuba-like protein, designated
XX  MFQ-114 polypeptide (I) and the polynucleotide (II) encoding it. (I),
XX  (II) and antibody (III) to (I) are useful for screening for identifying
XX  compounds that stimulate or inhibit the function or level of (I). (I) and
XX  (II) are useful for treating breast, prostatic, ovarian and pancreatic
XX  cancer, neoplastic disorders, neurodegenerative disease, brain stroke,
XX  cardiac and vascular disease, and angiogenesis, for inducing
XX  immunological response in a mammal, as vaccine, in disease diagnosis and
XX  in assays for screening agonistic or antagonistic compounds. (I) is
XX  useful for identifying membrane bound or soluble receptors, and in
XX  conventional low capacity screening methods and also in high-throughput
XX  screening (HTS) formats, and as immunogens to produce antibodies. (II) is
XX  useful as a diagnostic reagent for detecting mutations in the associated
XX  gene, in the recombinant production of (I), for chromosome localisation
XX  studies, in tissue expression studies, and in gene therapy. The
XX  present sequence represents the amino acid sequence of human Ajuba-like
XX  protein, MFQ-114.
XX  Sequence 538 AA;
XX  Query Match 60.3%; Score 44; DB 23; Length 538;
XX  Best Local Similarity 50.0%; Pred. No. 2.3e+02;
XX  Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY  2  GVCC-----GXXLCHXC 13
    | | | | |
Db  508  GCCCFPLDGHLLCHGC 523

RESULT 13
AAU31406
ID  AAU31406 standard; Protein; 93 AA.

```


XX AAU31406;
 AC 18-DEC-2001 (first entry)
 XX Novel human secreted protein #1897.
 DE Human; vaccination; gene therapy; nutritional supplement;
 XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 KW Homo sapiens.
 XX WO200179449-A2.
 XX 25-OCT-2001.
 PD 16-APR-2001; 2001WO-US08656.
 PF 18-APR-2000; 2000US-0552929.
 XX 26-JAN-2001; 2001US-0770160.
 PR (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-611725/70.
 DR Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy -
 PT Claim 20; Page 432; 765pp; English.
 PS The invention relates to novel human secreted polypeptides. The
 XX polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX Sequence 93 AA;
 SQ Query Match 58.9%; Score 43; DB 22; Length 93;
 Best Local Similarity 60.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 GVCCGXXLCH 11
 DB 65 GPCCGIFCH 74
 RESULT 14
 AAM79259
 ID AAM79259 standard; Protein; 766 AA.
 XX AAM79259;
 AC 06-NOV-2001 (first entry)
 XX Human protein SEQ ID NO 1921.
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW Human; cytokine; cell proliferation; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX nervous system disorder; arthritis; inflammation.
 OS Homo sapiens.
 XX WO200157190-A2.
 PN 09-AUG-2001.
 PD 05-FEB-2001; 2001WO-US04098.
 PF 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0683561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR N-PSDB; AAK52392.
 XX Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 PT Claim 20; Page 4320-4322; 6221pp; English.
 PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AA80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX Sequence 766 AA;
 SQ Query Match 58.9%; Score 43; DB 22; Length 766;
 Best Local Similarity 75.0%; Pred. NO. 4.le+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 CGGXXLCH 11
 DB 240 CCGNLLCH 247
 RESULT 15
 ABG05471
 ID ABG05471 standard; Protein; 1203 AA.
 XX ABG05471;
 AC 13-FEB-2002 (first entry)
 XX Novel human diagnostic protein #5462.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
 XX WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PA (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX DR N-P5DB; AAS69658.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX Claim 20; SEQ ID NO 35830; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 1203 AA;

Query Match 58.9%; Score 43; DB 22; Length 1203;
 Best Local Similarity 75.0%; Pred. No. 5.8e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
 ||| |||
 Db 737 CCGNILCH 744

Search completed: January 29, 2003, 09:30:58
 Job time : 14 secs